A Quantitative Analysis of Architectural and Engineering Procurement: Effects of Cost Inclusion on Procurement Outcomes Compared to Qualifications- Based Selection

By

Scholar Amirali Shalwani
B.Eng., NED University of Engineering and Technology, 2014

Submitted to the graduate degree program in Civil, Environmental & Architectural Engineering and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Master of Science.

Chairperson: Dr. Brian Lines, Ph.D.

Dr. Daniel Tran, Ph.D.

Michael A. Panethiere, P.E.

July 27th, 2017

The Thesis Committee for Amirali Shalwani

certifies that this is the approved version of the following thesis:

A Quantitative Analysis of Architectural and Engineering Procurement: Effects of Cost Inclusion on Procurement Outcomes Compared to Qualifications- Based Selection

Chairperson: Dr. Brian Lines, Ph.D.

July 27th, 2017

ABSTRACT

In 1972, the Brooks Act established Qualifications-Based Selection (QBS) at the federal level, requiring the U.S. Federal Government to procure architectural and engineering (A/E) consultants based upon their experience and capability rather than cost. Yet at the state, municipal, and local levels, and as well as in the private sector, best value procurement (BV) of A/E services has become prevalent as owners have utilized price evaluations with increasing regularity. A/E professionals have widely viewed this as a trend that carries threatening implications for their profession. In instances where owners insist on evaluating price, professional associations in the A/E community have responded by establishing guidelines for how price should be treated to ensure an effective, fair, and transparent selection process. The two most widespread recommendations are that (1) owners should limit the weight of pricing such that it does not dominate the selection outcome, and (2) a two-envelope system should be used to ensure the evaluation process is not unduly biased towards lowest price. Yet little research has investigated the application of these guidelines within the industry. This paper aims to address this gap by analyzing the characteristics of selected bidders from a dataset of 122 publically-procured A/E projects across North America, where the owners' evaluation process followed the guidelines of limited price weighting and a two-envelope system. Results showed that half the time, the consultant selected within the best value (BV) procurement system was the lowest bid. In the cases where the lowest bid was selected, results showed that the consultant was nearly always ranked within the top three qualifications scores among all consultants. In a quarter of the projects, the lowest bidder also had the highest qualifications. The range of owner evaluation scores for each competing consultant were also analyzed in order to identify which evaluation criteria achieved the greatest differentiation among consultants. Results indicated that cost submissions, schedule proposals, and interviews achieved greater differentiation than technical proposals, past performance or related experience criteria. Lastly, interrelationship of evaluation criteria showed that, no direct relationship existed between cost and other qualifications criteria. The results of this study can help owners in terms of how to optimize the inclusion and weighting of their evaluation criteria. Implications for A/E professionals include recommendations for how to strategize proposal plans to emphasize evaluation criteria that achieve greater differentiation in order to stand out from their competition. Furthermore, the approach of analyzing the evaluation scores for all bidders is novel, and can help researchers in future utilize similar approach.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor and committee chairperson, Dr. Brian Lines, for his continuous support and guidance throughout my thesis and master's program. I appreciate all the help and resources he has provided me, and for mentoring me in the best possible way. Working under him has helped me gain new skills and become better professionally.

I would also like to thank the committee members, Dr. Daniel Tran and Mr. Michael A. Panethiere, P.E., for their valuable support and time in reviewing my thesis and giving feedback when necessary.

I would also like to thank my parents who have always been my biggest supporters. It was because of their support and their faith in me that I have accomplished so much.

Lastly, I would like to thank my brothers, family, and friends for their continuous support and belief in me.

TABLE OF CONTENTS

ABSTRACT ... **iii**
ACKNOWLEDGEMENTS ... **v**
LIST OF FIGURES ... **ix**
LIST OF TABLES ... **x**
CHAPTER 1: INTRODUCTION ... **1**
1.1 BACKGROUND ... 1
1.2 RESEARCH OBJECTIVE AND METHODOLOGY ... 1
1.3 KEY FINDING ... 2
1.4 THESIS ORGANIZATION ... 2
CHAPTER 2: LITERATURE REVIEW ... **5**
2.1 PERCEPTION OF INDUSTRY PROFESSIONALS ... 5
2.2 CURRENT RESEARCH FINDING FOR A/E PROFESSIONALS ... 6
2.2.1 COMMODITIZATION OF DESIGN PROFESSIONALS ... 7
2.2.2 MULTICRITERIA SELECTION OF DESIGN PROFESSIONALS ... 8
2.2.3 PERFORMANCE OF DESIGN PROFESSIONALS ... 9
2.2.4 PRE-SELECTION CRITERIA OF DESIGN PROFESSIONALS ... 10
CHAPTER 3: RESEARCH OBJECTIVES AND HYPOTHESIS ... **13**
3.1 POINT OF DEPARTURE ... 13
3.2 OBJECTIVES ... 14
3.3 RESEARCH QUESTIONS ... 16
3.3.1 RESEARCH QUESTION 1 (RQ1) ... 16
3.3.2 RESEARCH QUESTION 2 (RQ2) ... 16
3.3.3 RESEARCH QUESTION 3 (RQ3) ... 16
3.3.4 RESEARCH QUESTION 4 (RQ4) ... 16
3.3.5 RESEARCH QUESTION 5 (RQ5) ... 17
3.4 HYPOTHESIS STATEMENTS ... 17
CHAPTER 4: METHODOLOGY ... **19**
4.1 RESEARCH FRAMEWORK ... 19
4.2 DATA COLLECTION ... 20
4.3 DEFINING VARIABLES ... 21
4.3.1 EVALUATION CRITERIA ... 21

4.3.2 PROJECT CHARACTERISTICS AND OTHER CLASSIFICATIONS 23
4.3.3 COEFFICIENT OF VARIATION (COV) 26
4.3.4 KRUSKAL-WALLIS H TEST VARIABLES 27
4.3.5 NORMALIZED COST AND SCHEDULE 28
4.4 EVALUATION PROCESS.. 29
4.4.1 TWO-ENVELOPE BV PROCUREMENT METHOD... 29
4.4.2 QUALIFICATIONS-BASED PROCUREMENT METHOD..................................... 33
4.5 DATA CHARACTERISTICS 36
CHAPTER 5: METHOD OF ANALYSIS... 37
5.1 DESCRIPTIVE ANALYSIS.. 37
5.2 KRUSKAL-WALLIS H TEST 38
5.3 SPEARMAN'S CORRELATION COEFFICIENT 39
CHAPTER 6: RESULTS 41
6.1 SELECTION OUTCOMES IN BEST VALUE PROCUREMENT OF A/E CONSULTANTS 41
6.2 EVALUATION CRITERIA EFFECTIVENESS IN DIFFERENTIATING CONSULTANT QUALIFICATIONS 44
6.2.1 KRUSKAL-WALLIS H TEST AMONGST ALL PROJECT CHARACTERISTIC.. 44
6.2.2 KRUSKAL-WALLIS H TEST AMONGST INDIVIDUAL PROJECT CHARACTERISTIC 48
6.2.3 OWNER PROVIDED BUDGET AND SCHEDULE... 53
6.3 RELATIONSHIP BETWEEN CONSULTANT QUALIFICATIONS AND BID COST . 54
CHAPTER 7: DISCUSSION, CONCLUSIONS, AND CONTRIBUTIONS 56
7.1 SELECTION OUTCOMES IN BEST VALUE PROCUREMENT OF A/E CONSULTANTS DISCUSSIONS 56
7.2 EVALUATION CRITERIA EFFECTIVENESS IN DIFFERENTIATING CONSULTANT QUALIFICATIONS 59
7.2.1 KRUSKAL-WALLIS H TEST AMONGST ALL PROJECT CHARACTERISTICS 59
7.2.2 KRUSKAL-WALLIS H TEST AMONGST INDIVIDUAL PROJECT CHARACTERISTICS 60
7.2.3 OWNER PROVIDED BUDGET AND SCHEDULE... 63
7.3 RELATIONSHIP BETWEEN CONSULTANT QUALIFICATIONS AND BID COST . 64
7.4 CONCLUSIONS 64
7.5 CONTRIBUTIONS.. 66

7.5.1 CONTRIBUTIONS TO THE BODY OF KNOWLEDGE 66
7.5.2 CONTRIBUTIONS TO INDUSTRY PRACTITIONERS 67
7.6 LIMITATION AND RECOMMENDATIONS FOR FUTURE RESEARCH 69
REFERENCES 70
APPENDIX A – BOX PLOTS AND SCATTER PLOTS 73
APPENDIX B – COST-QUALIFICATIONS MATRIX AND POST HOC TESTS 92

LIST OF FIGURES

Figure 1. Stanford's CIFE Horseshoe Research Framework 20

Figure 2. Evaluation Process for Two-envelope Best Value Procurement Method 32

Figure 3. Evaluation Process for Qualifications-Based Procurement Method 35

Figure 4. Ranking of Selected Bidders for All Evaluation Criteria 44

Figure 5. Venn diagram for Selected BQ and LB Firms 57

LIST OF TABLES

Table 1. Journals Reviewed during Literature Review 7

Table 2. Number of Projects for Each Project Characteristics' Sub-Factor 24

Table 3. Descriptive Analysis of Evaluation Criteria Weight 36

Table 4. Characteristics of Selected Bidders 41

Table 5. Cost-Qualification Matrix for Selected Bidders 42

Table 6. Differential between Selected Bidders and Competing Proposals 43

Table 7. Descriptive Statistics for Coefficient of Variation 45

Table 8. KW-H Test for All Evaluation Criteria 45

Table 9. Post-Hoc Testing for All Evaluation Criteria Combination 46

Table 10. KW-H Test and Classification for Project Characteristics 47

Table 11. KW-H Test Evaluation Criteria Differentiation Based on Scope of Work 48

Table 12. KW-H Test Evaluation Criteria Differentiation Based on Project Type 49

Table 13. KW-H Test Evaluation Criteria Differentiation Based on Procurement Type 50

Table 14. KW-H Test Evaluation Criteria Differentiation Based on Prime Vendor 50

Table 15. KW-H Test for Evaluation Criteria Differentiation Based on Size of Project 51

Table 16. Post-Hoc Test for Size of Project 51

Table 17. KW-H Test between All Criteria and Project Complexity 52

Table 18. Post-Hoc Test for Project Complexity 53

Table 19. KW-H Test for Owner-Provided Budget and Schedule against Median Cost and Schedule COV 54

Table 20. Spearman's Correlation Coefficient for All Evaluation Criteria Combination 55

CHAPTER 1: INTRODUCTION

1.1 BACKGROUND

Qualifications-Based Selection (QBS) has been the predominant procurement method for architectural and engineering (A/E) services for more than fifty years (Chinowsky and Kingsley 2009). The Brooks Act, which was adopted by Congress in 1972, requires all federal projects to procure architectural and engineering (A/E) services on the basis of qualifications and then negotiate a fair and reasonable cost with the selected bidder. According to the American Council of Engineering Companies (ACEC), forty-six out of fifty states have emulated the federal government by adopting their own QBS laws, oftentimes called "mini-Brooks" acts. Although almost all states have what is considered a "mini-Brooks" act at the core of their procurement law, the extent to which this core has exceptions varies from state to state. As states continue to integrate alternative delivery strategies (such as design-build, construction manager, and integrated project delivery), they are presented with the opportunity to adjust procurement law provisions, which has tended to weaken QBS procurement requirements (Chinowsky and Kingsley 2009). For public entities that are further removed from the federal government, such as at the municipal and local levels, the departure from QBS requirements in favor of a price-focused procurement process becomes even more noticeable.

1.2 RESEARCH OBJECTIVE AND METHODOLOGY

A/E professionals have long been in disagreement with clients over the inclusion of cost proposals during the procurement process and viewed it as a threat to the profession. Therefore, it became important to analyze the affect that cost criteria has on the procurement outcomes for A/E selections. Selected bidders' characteristics were analyzed using descriptive statistics, while Coefficient of Variation (COV) was utilized to find what level of differentiation existed between all evaluation criteria. Determining the relationship of cost criterion with other qualifications criteria was also important to corroborate design professionals' claim that more cost brings more expertise. Descriptive analysis, Kruskal-Wallis H test, and Spearman's correlation coefficient were statistical measures used to analyze 122 projects.

1.3 KEY FINDING

An analysis of the owner evaluation scores in the data sample revealed that almost 48% of the time, the selected bidder was the lowest bid compared to 54% of the time when it was the most qualified. The results also showed that a wide range of cost proposals were usually submitted when cost was included as an evaluation criterion. Schedule and consultant team interviews achieved a similar level of differentiation. Technical proposals achieved a moderate level of differentiation, while past performance information and related experience achieved the least. Cost proposal and schedule proposal also achieved a greater level of differentiation when the owner provided a design budget and schedule. Lastly, cost was found to have no relationship with any of the qualification criteria.

1.4 THESIS ORGANIZATION

The literature review in Chapter 2 presents the plan of reviewing different journal and industry professionals' advocacy on design professionals' procurement. The literature review was divided into two parts, industry professionals' perception and current research finding on design professionals' procurement process. All relevant papers to this thesis were categorized into four main areas, commoditization, multi-criteria selection, performance, and pre-selection criteria of design professionals.

Chapter 3 consists of points of departure from previous research, objectives of the study, research questions in order to reach the objectives, and lastly all hypothesis statements developed which helped in reaching conclusions with statistically proven tests.

Chapter 4 describes the methodology adopted to accomplish the goals for the thesis. One hundred and twenty-two A/E procurement projects from various industry sectors were analyzed using descriptive and inferential tests. This chapter also defines the variables used by descriptive and inferential analysis, such as different evaluation criteria and project characteristics, and the procurement evaluation process for best value and qualifications-based systems. At the end, the chapter illustrates the descriptive statistics of the six evaluation criteria weights used across all the projects analyzed.

Chapter 5 consists of various methods of analysis employed to analyze the data. Descriptive analysis, mainly the measure of central tendency, frequency distribution tables, and matrix representations, was used in order to describe the characteristics of selected bidders. Whereas, Kruskal-Wallis H test and Spearman's correlation coefficient were used as inferential tests for hypothesis statements established in the previous section.

Chapter 6 of the thesis consists of results and findings for all statistical measures described in the previous section.

Chapter 7 discusses all the results and findings, as well as the conclusions to the hypothesis statements established in earlier sections. Each result and finding was linked back to the research questions and how it helped in answering it. Lastly, the author provides the contribution of the research towards academics and industry while giving recommendations and improvements for future research.

Part of this thesis, which included the characteristics of selected consultant and descriptive for the six evaluation criteria's Coefficient of Variation (COV), was published and presented by the author in AEI conference held in April of 2017 at Oklahoma City. The citation for the conference paper is as follow,

- Lines, B, C. and Shalwani, A, S. (2017) "Best Value Procurement of Architectural and Engineering Services: Selection Characteristics and the Relative Influence of Various Evaluation Criteria." AEI Conference 2017.

CHAPTER 2: LITERATURE REVIEW

Chapter 2 provides an overview of the general perception amongst the industry professional for the use of QBS as a procurement method for A/E firms. Unlike owners, Industry professionals oppose the use of cost as an evaluation criteria, and endorse the use of QBS. This chapter also conducts a thorough literature review of studies focused on various topics related to the procurement of A/E services.

2.1 PERCEPTION OF INDUSTRY PROFESSIONALS

The official position of professional associations within the A/E industry has been to promote the usage of QBS methods. QBS consists of selecting design professionals purely on qualifications, while the cost is negotiated once the consultant is selected. When the owners use BV procurement methods, consultants strive to limit the evaluation weight allocated for cost criterion. The American Institute of Architects (AIA) strongly supports QBS for procuring design professionals for public projects, based on their reasoning that "QBS provides the owners with a selection process that is not only straightforward and easy to implement, but is objective and fair" (AIA 2011). Similarly, the American Council of Engineering Companies (ACEC) has highlighted QBS as the "cornerstone" of their procurement policies (ACEC 2006). The American Public Works Association (APWA) believes that procuring A/E services using qualifications rather than cost "fosters greater creativity and flexibility, and minimizes the potential for disputes and litigation" (APWA 2008). The National Society of Professional Engineers (NSPE) "strongly" supports the Brooks Act and recommends that the procurement of design professionals in all sectors should utilize QBS methods (NSPE 2016). The American Water Works Association (AWWA) specifically encourages water utilities to procure A/E services on the basis of QBS (AWWA 2016).

2.2 CURRENT RESEARCH FINDING FOR A/E PROFESSIONALS

Literature review of the last ten years from renowned journals was conducted. As shown in Table 1, research studies on A/E procurement were found to be scarce, and even those found relevant mostly focused on commoditization of design professionals, multi-criteria decision making procurement, pre-qualification of design professionals, and performance of design professionals. None of these research papers analyzed the results of BV or QBS procurement method. Because of such a dearth of previous research in this field, the search was expanded. As procurement of design-build teams also incorporates the selection of a design firm with a contractor, it was considered the next closest thing to procurement of design professionals. Numerous research studies have been conducted for design-build BV procurement, but again most of this research was focused on evaluation criteria selection for BV procurement of DB firms, performance of different types of procurement methods, and DB BV procurement models rather than an actual evaluation of the BV procurement process. Furthermore, a search using key words such as ‘best value,’ ‘qualifications-based system,’ ‘architecture,’ ‘engineering,’ ‘procurement,’ etc. was also run on KU libraries, University of Kansas’s official literary search engine, for different journals in order to find as many relevant research papers as possible. Studies found most relevant for design professionals investigated a narrow range of research objectives, which can be categorized into four main areas, commoditization, multi-criteria selection, performance, and pre-selection criteria of design professionals.

Table 1. Journals Reviewed during Literature Review

S.no	Name of Journal	Number of Relevant A/E Procurement Papers Found
1	Journal of Management in Engineering	3
2	International Journal of Project Management	2
3	Construction Management and Economics	2
4	Engineering Construction and Architectural Management	2
5	Journal of Professional Issues in Engineering Education and practice	2
6	Journal of Construction Engineering and Management	1
7	Journal of Architectural Engineering	0
8	Architectural, Engineering and Design Management	0
9	International Journal of Procurement Management	0
10	Canadian Journal of Civil Engineering	0
11	Others	6

2.2.1 COMMODITIZATION OF DESIGN PROFESSIONALS

The major theme within the literature review was the commoditization of design professionals, which refers to the increasing trend towards the usage of cost based procurement selection processes. Hampton (1994) describes the owners' reservation to procuring design professionals using QBS. They claim QBS encourages the design professionals to bid high, while the professionals think the contrary, stating the owners are trying to commoditize the profession of design by incorporating cost during selection. Furthermore, the subjective procurement process has made it difficult for small, minority, and women-based firms to compete in QBS, which again thwarts the owners from using QBS. Moreover, Horns and Jenkins' (2011) response to the owners' perception of treating design services as a commodity was that services provided by design professionals cannot be readily quantified or defined. Additionally, they found qualified design professionals charge more for the expertise and qualification they bring into the project. This paper further concluded that outsourcing work to offshore engineering firms has also led to design professions being commoditized as the owners look for firms that can do the job the cheapest, while potentially compromising on the quality of work. Commoditization of design professionals has led to much bigger problems. Parks (2006) concludes that due to a 7% increase in salaries of design professionals compared to 35% to 40% in other professions, less students are pursuing this field. As design professionals are still paid according to Time-based Compensation (TBC), Parks suggested that design professionals should be paid by Value- based Compensation (VBC) in order to compensate professionals for the quality of work rather than the quantity of work done.

2.2.2 MULTICRITERIA SELECTION OF DESIGN PROFESSIONALS

The second important area in the literature review was the selection of design professionals using multi-criteria selection techniques. Research conducted in this area was mainly focused on the different multi-criteria procurement models developed by different researchers for selecting the design professionals. The criteria considered for selection are based on subjective analyses from industry professionals or from widely used industry practices. In a study based in Hong Kong, Cheung *et al.* (2002) developed a multi-criteria selection based computer model called "Architectural Consultant Selection System" (ACSS) to assist the owners in selecting the best consultant for their project. Nguyen and Shehab (2008) also developed a multi-criteria selection model using fuzzy set theory to select the most qualified A/E. In addition to the tender cost, the model also used qualifications-based criteria such as financial soundness, experience, expertise, availability, and compatibility of personality. Ling (2003) developed a model for selecting architects by project managers using theory of firm (price factor), theory of task performance, theory of contextual performance, and theory of embeddedness (network factor). A total of forty evaluation criteria were shortlisted, out of which thirty-four criteria were found to be important when surveys from 200 project managers and 120 architects were conducted. They concluded that the theory of contextual performance and task performance were found to be important, while the price factor was not considered to be an important criterion when procuring architects. Sullivan *et* *al.* (2010) developed a BV Performance Information Procurement System (PIPS) business model which uses Past Performance Information (PPI), Risk Assessment and Value Added (RAVA), and interview as its evaluation criteria when procuring design professionals. The business model was implemented for many public projects and the results were very encouraging. The owner was able to reduce the procurement time, with fewer change orders, reduced project delays, and higher customer satisfaction. Moreover, all projects analyzed in this thesis were procured using the same BV PIPS business model.

2.2.3 PERFORMANCE OF DESIGN PROFESSIONALS

This section discusses the performance of design professionals. Different researchers have analyzed the performance of design professionals for different delivery methods and procurement types for various cost, schedule, and quality metrics. Ling (2002) identified attributes that affect an A/E's performance in Design-Build delivery projects. The study developed a model based on twenty-four criteria that helped predict the performance of the consultant. The study also showed that 86% of A/E's performance can be predicted using just three criteria: problem-solving ability and project approach, speed in producing design drawings, and enthusiasm in tackling a difficult assignment. Ng and Chow (2004a) also developed a multi-criteria model for evaluating the performance of the engineering consultant. The model was devised using consultant performance evaluation framework, which resulted in "achievement of objectives and targets," "quality of bid documents," "compliance to clients' requirements," "compliance to legislative requirements," and "identification of clients' requirements and project objectives" being the most important criteria. Oyedele and Tham (2007) conducted a survey in which seventy-one public and private clients were asked to assess architect performance in twenty-eight criteria. The results showed that architects did not perform well in twenty criteria, which were classified into six categories:

management skills and ability, buildability of design, design quality, project communication, project integration, and client factor. The paper also concluded that architect performance was 74% or less for twenty criteria and therefore needed to be evaluated further. With many clients asking for past performance as an evaluation criterion, Chow and Ng (2007a) identified a unified set of Consultant Performance Evaluations (CPEs) to evaluate the performance of design professionals. A survey questionnaire was sent out to 100 industry professionals to rate eight CPEs consisting of eighteen quantifiable identifiers. Thirty-two valid responses were received and the coefficient of variation for each CPE was calculated, which resulted in large values, showing a higher differentiation of scores. In the follow up, Chow and Ng (2007b) developed a fuzzy gap model to score A/E firms on the basis of their performance and initial expectation of the client. The model can help owners make quantifiable performance scores, which can further assist clients who asked the AE firms for past performance information during the procurement stage.

2.2.4 PRE-SELECTION CRITERIA OF DESIGN PROFESSIONALS

The fourth and last research objective found in the literature was pre-selection criteria for design professionals. This section discusses the different models developed by researchers for pre-selecting design professionals based on criteria established though subjective analysis from industry professionals. With respect to the other three categories, not many papers were found for this category. Through an extensive survey, Ng and Chow (2004b) identified twenty-eight pre-selection criteria that can be classified into four major criteria: technical capability, management capability, financial capability, and quality assurance and quality control (QA/QC). The survey results also showed that technical capability was considered the most crucial, while the QA/QC criteria was the least important. Feldmann *et al.* (2008) developed a multiple regression model to estimate the cost for A/E services for university projects. Fifty-eight factors were identified from

165 building projects across fifty-two research universities. Only nine factors were found to be significant, and project complexity was not one of them. Gross square feet of the building, physical plant expenditure, institutional endowment in 2004, institutional control, and region as the independent variable accounted for 45.9% of the variance in the A/E cost. This model can help institutional management predict the cost for design professionals for university projects.

The most well-known and comprehensive research study on QBS was conducted by Chinowsky and Kingsley (2009). The objective of the study was to provide a quantitative analysis of projects procured using QBS and analyze its benefits when compared to other non-QBS projects. The study had a sample size of forty-two projects that included full project data, had a public project representation of 95%, and represented projects from 29 states from the United States. However, only 78% of these projects were procured using QBS and 10% using BV, which aggregated to a total of thirty-six projects being procured using QBS or BV, while the remaining five were either procured using low bid or sole source. The results showed that QBS procured projects had a construction cost growth and schedule growth of 3% and 8.7%, respectively, based upon categorical survey results, which were inferred to be below industry averages. Furthermore, a questionnaire was developed for the clients and design professionals regarding the attributes of QBS procurement. The results of the survey showed that the owners preferred using the QBS system when the design complexity was high and the risks were greatly reduced. Additionally, the owners rated the project success of QBS projects to be high or very high 93% of the time. The survey also showed that QBS cultivated the sense of embeddedness and trust between the owner and the client. However, the study’s dataset was deemed to include too few responses to support inferential statistical tests.

In the other most relevant study for the selection of design professionals, Christodoulou *et al.* (2004) analyzed 162 A/E projects from New York City, which were procured using a two-envelope BV procurement method. The city claimed to have saved significant tax payer dollars by injecting cost competition to A/E selections. However, the study noted that by considering cost as an evaluation criterion, the city encouraged consultants to bid as low as possible to get projects, thereby reducing project quality. In order to document the performance results of BV procurement, the city compiled a large dataset of recent projects and calculated the cost performance results, which showed a significant amount of money was saved when compared to procuring using QBS. However, Christodolou *et al.* found the city's dataset to be highly skewed, as it contained a substantial number of non-professional service projects. In reality, the savings achieved by the city should be valued at 1.67% of their claimed savings of $892 million. The study further concluded that the city's report failed to report the impact design service procurement had on the final cost of the project.

Even with an extensive literature review from several renowned journals from the past ten years and with a specific keyword search, only two papers were found to be most relevant to this thesis's research area. Still, these two papers had quite a few points of departure in terms of analysis and sample size.

CHAPTER 3: RESEARCH OBJECTIVES AND HYPOTHESIS

Chapter 3 reviews the point of departure from the existing literature along with the research objectives for this thesis. Furthermore, five research questions and five hypothesis statements were developed in order to meet those objectives.

3.1 POINT OF DEPARTURE

The existing literature related to procurement of A/E professionals has been primarily limited to a small sample sizes. The most prominent study of QBS, by Chinowsky and Kingsley (2009), analyzed only forty-two projects for their performance in cost, schedule, and quality, with only 95% representation of public projects. Among these projects, only thirty-three were procured via QBS and four were procured using best value (BV). However, for these BV projects the study did not specify whether the recommended evaluation guidelines were followed. The thesis fills in the gap by analyzing a relatively larger sample size, with 102 two-envelope BV and twenty QBS projects. All the projects in the data sample were public projects. Due to a dearth of research, this study was the most relevant research to the thesis topic.

Other important research to the thesis was limited to a specific geographic location. Christodolou *et al.* (2004) examined the authenticity of the New York mayor's office claim of saving tax payers' money when projects were procured using competitive bidding, two-enveloped BV, rather than QBS. The study analyzed 162 projects and rejected the mayor's office's claim stating most of the projects in the sample contained non-professional services (around 118). Also, no change order cost for projects were accounted for when calculating the actual cost of the project, resulting in fewer savings. This thesis fills in the gap by analyzing relatively larger sample size, which furthermore consisted of evaluation scores from 804 proposing consultants, from across North America at various levels of the public sector.

Other research was focused on multi-criteria analysis, in which only the content analysis of the evaluation weight was conducted from the various RFQ/RFP and based on factor analysis models developed for design professionals' selection. This thesis analyzes the procurement process with actual evaluation scores assigned by the evaluation team during the procurement process. Furthermore, the six evaluation criteria were thoroughly analyzed for ranks, scores, and inter-comparison. Inferential testing was also conducted to determine the level of differentiation amongst criteria, while cost criterion was examined for relationships with other evaluation criteria.

3.2 OBJECTIVES

Inclusion of cost in the procurement process has always been a topic of debate between many design professionals and the owners. Design professionals claim that design services should not be procured on the basis of cost, as each firm brings in its own expertise and qualifications, which cannot be quantified. On the other hand, the owners still use the traditional way of procuring, lowest bid, stating it is transparent, saves tax-payers money, and an easy procurement process. Brook's Act, which was passed by Congress in 1972, made it compulsory for federal projects to procure design services using only qualifications-based systems. The Act does not allow public project owners to ask design firms for cost until they are selected purely on qualifications, and later cost is to be negotiated. However, the Act was interpreted differently by different states, cities, municipalities, and at other government institution levels, which led to the development of the BV procurement model, which incorporates both cost and qualifications.

Design professionals are generally opposed to a substantial introduction of cost criteria within owner procurement methods**.** Christodolou *et al.* (2004) stated that when low bidding is used, the owner essentially commoditizes design professionals. In some cases, when cost was used as a dominant evaluation factor, the effect was that consultants are essentially being selected on the basis of lowest bid (Chinowsky and Kingsley 2009). Internationally, in order to mediate the low bid selection criteria of consultants, the World Bank uses the best value procurement method and states that if design professionals are chosen from a weighted quality and cost scores, cost should not exceed thirty weighted points out of a total score of 100 weighted points (Fleenor and Hall 2002). Wardani *et al.* (2006) even assumed projects to be QBS if the cost weight was 50% or less. Even when owners operate within those guidelines, very little research has documented the selection outcomes and the influence cost has on the procurement process. Therefore, it becomes important to analyze the effect of cost on the procurement process, even after limiting the weight to 30% or less.

The purpose of using the two-envelope BV procurement process was to select the firm which was the best value, cost and qualifications criteria combined. It was important for the owner to choose the criteria which could help them procure the best value firm for their projects. Therefore, the other objective of the study was to find which of the six evaluation criteria gave the greatest differentiation in scores. Also, with cost being a major factor when procuring a design professional, this thesis also investigated if the owner-provided design budget and schedule could help the owner select firms equal or very close to the design budget and schedule.

Design professionals further claim that when they quote high costs, it is because of the qualifications and expertise they bring to the project. Furthermore, they criticize the LB procurement method as it gives an extra edge to firms who bid lowest, while delivering lower quality projects. Hence, the last objective of the research was to determine if there exists a relationship between the cost and other qualifications criteria of the firm which can corroborate to the design professionals' claim that higher cost proposals bring higher qualifications.

3.3 RESEARCH QUESTIONS

With the increasing use of BV and QBS procurement methods for hiring design professionals, it becomes important to analyze both the procurement method for its effectivity and performance of bidders in the process. This research analyzes selected bidders' performance and level of differentiation achieved by the six common evaluation criteria, and if any relationship exists between cost and other qualifications criteria.

The research focuses on five major questions. Statistical analysis, both descriptive and inferential statistics, has been used in order to answer these questions.

3.3.1 RESEARCH QUESTION 1 (RQ1)

What are the selection outcomes of the two-envelope best value procurement processes for architectural and engineering services? Does the inclusion of cost as an evaluation criteria have a disproportionate effect on selection outcomes, such that the owners trend towards the selection of lowest bid or non-highest qualifications?

3.3.2 RESEARCH QUESTION 2 (RQ2)

Which evaluation criteria achieves the greatest level of differentiation between competing A/E proposals?

3.3.3 RESEARCH QUESTION 3 (RQ3)

Do evaluation criteria achieve greater differentiation depending upon various project characteristics (i.e. scope definition, project type, procurement type, prime vendor, project size, project complexity)?

3.3.4 RESEARCH QUESTION 4 (RQ4)

Does the publication of a project budget and/or schedule within the owners' RFP have an effect on the submitted cost and schedule proposals by A/E firms?

3.3.5 RESEARCH QUESTION 5 (RQ5)

Does the selection of greater qualifications among A/E firms correspond with higher cost proposals?

RQ1 employs basic descriptive statistics, while RQ2, RQ3, and RQ4 require inferential testing using the Kruskal-Wallis H test. Spearman's correlation coefficient would be used for RQ5.

3.4 HYPOTHESIS STATEMENTS

RQ1, which analyzes the characteristics of the selected bidder, would use differential statistics, hence no hypothesis statement was developed for reaching a conclusion. Hypothesis statements $H1$ and $H2$ were developed in order to answer RQ2 and RQ3, respectively, which determines the level of differentiation each evaluation criteria achieves. Hypothesis statements $H3$ and $H4$ answer RQ4, which analyzes the effect of owner provided budget and schedule information in the RFP on the cost and schedule proposals. The last hypothesis statement, $H5$, was used to answer RQ5 regarding the relationship between cost and other qualifications criteria. Furthermore, a lot of inferential tests were conducted for RQ2 and RQ3, and for each test conducted, a hypothesis statement was developed. $H1$ and $H2$ are overall hypotheses, which are dependent on the results of all hypotheses developed while conducting tests.

$H1_0$: Different evaluation criteria will not result in various levels of differentiation between competing A/E firms.

$H1_1$: Different evaluation criteria will result in various levels of differentiation between competing A/E firms.

$H2_0$: The evaluation criteria will not achieve various levels of differentiation based upon separate project characteristics.

$H2_1$: The evaluation criteria will achieve various levels of differentiation based upon separate project characteristics.

$H3_0$: The project budget provided will not result in lesser variation of cost proposals (lower values of cost Coefficient of Variation (COV)).

$H3_1$: The project budget provided will result in lesser variation of cost proposals (lower values of cost Coefficient of Variation (COV)).

$H4_0$: The provided schedule will not result in lesser variation of schedule proposals (lower values of schedule days Coefficient of Variation (COV)).

$H4_1$: The provided schedule will result in lesser variation of schedule proposals (lower values of schedule days Coefficient of Variation (COV)).

$H5_0$: No directly proportional relationship exists between the cost criteria and other qualifications criteria.

$H5_1$: A directly proportional relationship exists between the cost criteria and other qualifications criteria.

CHAPTER 4: METHODOLOGY

Chapter 4 presents the methodology used for this thesis. A research framework was developed which briefly describes the road map of the thesis; including the research need, data collection and analysis, and recommendation to industry practitioners as well as future researchers. Furthermore, this chapter describes the two-envelope BV and QBS procurement process that was used by the projects within the data sample. Definition of variables is also provided for the six evaluation criteria and different project characteristics used to conduct analysis. A descriptive analysis for the weights used by the owners for the six evaluation criteria was provided at the end of the chapter.

4.1 RESEARCH FRAMEWORK

Figure 1 shows the research framework for the thesis using Stanford's CIFE research framework. The framework starts by identifying the problem, which for the thesis was the inclusion of cost as an evaluation criteria and at times the only criteria, leading to commoditization of A/E professionals. Not much research has been conducted when analyzing the effect of cost in the procurement of A/E professionals. The most relevant research conducted was either small in sample or limited to a much smaller geographic location, and still they did not analyze the A/E procurement process. The thesis fills in the gap by analyzing 122 A/E projects from state, city, local, municipality levels, and university, for characteristics of the selected bidder, the level of differentiation achieved by different evaluation criteria, and the relationships of cost with other qualifications criteria.

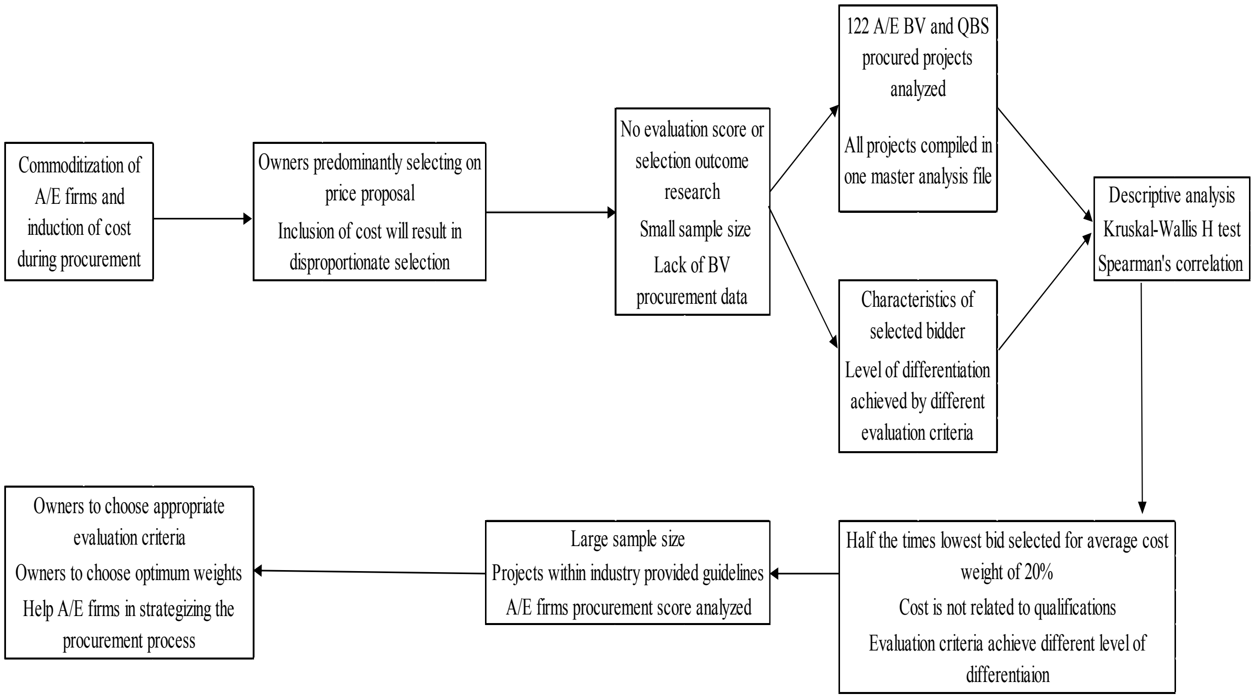

Figure 1. Stanford's CIFE Horseshoe Research Framework

The results show that in the A/E procurement process, almost half the time, the lowest bidder was awarded the project. Furthermore, cost proposals and schedule proposals achieved the highest level of differentiation in scores and cost was found to have no relationship with any of the qualifications criteria. The last part of the framework enumerates the contribution of the research. The research can help owners chose appropriate evaluation criteria with optimum weight in order to procure the best A/E firm. It can also help A/E professionals strategize their bidding process, while concentrating on evaluation criteria that can increase their chances of being selected. A detailed discussion for each of the topics in the framework is provided in the following sections.

4.2 DATA COLLECTION

The data sample was collected from various A/E projects at the state, city, and municipal levels, along with institutions of higher education, across North America. The sample size consisted of 122 A/E projects, out of which 102 were procured via a two-envelope BV method and the remaining twenty were procured using QBS.

All projects in the data sample used virtually identical evaluation criteria with similar weighting schemes. The main evaluation criteria included cost, schedule proposal, written technical proposals, past performance of the firm and their project team, consultant team interviews, and related experience (RE) of the firm. Within the two-envelope BV procured projects, cost was limited at 10-30% of the total evaluation weight. As this is in accordance with recommendations of A/E professional associations, the data sample was deemed to be an important contribution to the literature.

4.3 DEFINING VARIABLES

4.3.1 EVALUATION CRITERIA

Evaluation criteria, as the name implies, were criteria that the owners utilized during the procurement process in order to determine the best firm for the project. This research focused on the six common evaluation criteria: cost proposal, schedule proposal, technical proposals, past performance information, interview, and related experience. The descriptive statistics for the evaluation criteria weight were summarized in Table 3.

Cost. Cost criterion was the total design cost of the project that the bidder quoted in their bid proposal. For qualifications-based systems, no weight was assigned to the cost component of the proposal; project cost was later negotiated with the selected bidder. Cost criterion was evaluated based on the lowest bid. The bidder with the lowest cost was assigned the highest weight; all other bidders were scored using the inverse proportions method. Cost criterion was generally assigned a mean weight of 20% from an overall 100%.

Schedule. (Schedule or schedule proposal are used interchangeably in this thesis). This evaluation criterion asks for an anticipated project duration from the proposing bidder. Like cost, bidders with the lowest schedule were assigned the highest score, while all other bidders were scored using the inverse proportion method. The mean weight for schedule criterion was 14%.

Technical Proposals. Technical Proposals (TP) were comprised of three different proposals: project capability (PC), risk assessment plan (RA), and value assessment plan (VA). Some projects evaluated all proposals together and termed it as risk assessment and value added (RAVA) or project assessment plan (PA). When treated separately, all these criteria were weighted between 10-15% each. When combined, the mean weight was 34% out of an overall 100%. The standard format consists of writing a synopsis of the execution plan and enumerating all project specific risks and scope alternations/innovations.

Past Performance Information. Past Performance Information (PPI) was a reflection of the bidders' previous experience in similar projects with past clients (or ongoing projects). Within this study, past performance was evaluated as a survey from each bidder's previous client(s). Past performance scores were based on responses from the client(s) regarding the bidders' ability to manage cost, schedule, and quality. They were also scored on their communication, professionalism, ability to identify and minimize risk, and the clients' overall satisfaction with the bidders' performance. Past performance surveys were scored on a 1 to 10 scale, with 1 being the lowest score and 10, the highest. Like cost criterion, PPI also had a mean weight of 20%.

Consultant Team Interviews. (Referred to as interview from here on.) Interviews were conducted with key personnel from each shortlisted bidder. Interviews were mostly conducted with the top three to five bidders depending on the company's policy or the owners' will. The data set had a maximum of seven bidders being interviewed with a total number of nine proposals, while the data set also encountered a project in which three bidders were interviewed with a total of twenty-six proposals. Interviews were generally assigned the most weight in the evaluation process with a mean weight of 35%.

Related Experience. Related Experience (RE) criterion required bidders to submit a summary of their past projects related to the bidding project. Usually bidders provided summaries of projects which were either famous, larger in size, or for big companies. RE was assigned weight in between 5% to 30% of the total project weight.

Other Criteria. Some owners utilized other evaluation criteria in their procurement process, such as Women and Minority Business Requirements (WMBE) or team qualification. These criteria were not considered within the analysis due to the fact that they were typically assigned very low weight, between 5% to 10%, and were inconsistently applied across the dataset.

4.3.2 PROJECT CHARACTERISTICS AND OTHER CLASSIFICATIONS

In order to analyze the sample in more detail, six project characteristics were formed. The sample size was then further categorized into various sub-factors for each project characteristics. Table 2 below shows the distribution of all project characteristics' sub-factors across 122 projects in the data sample. Following is the list of project characteristics with a detailed description of each of them:

1) Scope of Work
2) Project Type
3) Procurement Type
4) Prime Vendor
5) Size of Project
6) Project Complexity

Table 2. Number of Projects for Each Project Characteristics' Sub-Factor

Project characteristics	Sub-factor	N
Scope of work	Pre-Design	19
	Detailed Design	103
Project type	Vertical	97
	Horizontal	25
Prime vendor	Architectural	77
	Engineering	45
Procurement method	BV	102
	QBS	20
Size of project	Small	37
	Medium	30
	Large	35
Project complexity	Low complexity	35
	Moderate complexity	31
	High complexity	56

Scope of Work. Scope of work was defined by the design phase for which the A/E firm is being procured. Scope of work had two sub-factors, pre-design and detailed-design. In pre-design projects, the owners procured A/E services for either a feasibility report, Design Concept Report (DCR), or any specialized engineering assessment. Pre-design projects were only limited to pre-design study of the project. Whereas, in Detailed-design projects, the owners procured A/E firms to either conduct a pre-design study and develop the design through construction documents, or the pre-design was provided and the selected bidders were required to start from schematic design and design through the construction documents.

Project Type. Project type was categorized into two sub-factors, vertical and horizontal construction projects. All the construction projects can be divided into these two classifications, therefore, the analysis based on vertical and horizontal projects was found relevant. Vertical projects included waste station treatment plants, all university buildings (halls, dining rooms, student housing, libraries, unions, etc.), and building systems. Horizontal projects were limited to roads, pavements, culverts and bridges, parks, and non-building pipeline systems.

Procurement Type. Procurement type was divided into QBS and BV procurement method. The difference between the two procurement methods was the inclusion of cost proposal in BV procurement, whereas no cost proposal was required in the QBS system. All projects with cost proposals were categorized as BV, while those without were categorized as QBS.

Prime Vendor. The prime vendor for a project was the business entity that the owner contracted with. They were also the prime consultant for the project. The data encountered a lot of projects in which the engineering firm was contracted as the prime vendor with an architectural firm as a sub-contractor, and vice versa. When an architectural firm was the prime vendor, generally the lead architect or lead architect project manager was interviewed, while when an engineering firm was the prime vendor, a lead engineer or a lead engineer project manager was interviewed. At times, it was explicitly stated in the RFQ/RFP who the owner wanted to contract with.

Size of Projects. All projects in the sample size were classified into small, medium, and large projects based on their average proposed bids. Projects with proposed average bids less than $26,000 were categorized as small projects. Those with proposed average bids ranging in between $26,000 to $99,000 were medium projects, while projects greater than $99,000 were categorized as large projects.

Project Complexity. All projects in the sample size were categorized as low, moderate, or high complexity. These categories were taken from Feldmann *et al.* (2008), who categorized projects into five different categories. The first category, or one, was least complex, and the last category, or five, was the most complex. Due to the sample size being too small in a few of the categories as well as a lot of projects in the sample not being categorized initially by Feldmann *et al.* (2008), a few assumptions were made when categorizing such projects. Categories one and two, the least complex projects, were combined to form low complexity projects for this research, while the two most complex categories, four and five, were combined to form high complexity projects. Category three was used for moderate complexity projects. Low complexity projects included parking garages, utility structures, apartments, and office buildings, while moderate complexity projects consisted of schools – elementary, junior high, and gymnasiums. Lastly, high complexity projects consisted of hospitals, science buildings, broadcast facilities, museums, power plants, and historic preservation facilities.

One of the objectives of the study was to analyze the effect of owner-provided project budget and schedule information on submitted cost and schedule proposals. The sample was further categorized into two other groups based on the owner-provided information in the RFQ/RFP:

1) Owner-Provided Budget
2) Owner-Provided Schedule

Owner-Provided Budget. This grouping was done based on projects where the owner provided some kind of budget information in the RFQ/RFP. When the owner provided either a design or construction budget or both, it was considered to be a 'yes,' whereas if no information was provided for the project budget, it was considered a 'no.'

Owner-Provided Schedule. This group was also divided on the basis that when the owner provided an anticipated construction start date or schedule duration of the project, it was considered to be a 'yes,' otherwise it was a 'no.'

4.3.3 COEFFICIENT OF VARIATION (COV)

Coefficient of Variation (COV) was employed to determine which of the six evaluation criteria achieved the greatest differentiation. COV is a measure of spread that describes the variability relative to the mean. To elaborate further, COV for each evaluation criteria describes the variation in scores for that particular evaluation criteria. For instance, higher values for COV denotes a larger variation in scores (higher range), while lower values of COV shows a lesser variation in scores (lower range). Mathematically, COV is the ratio of standard deviation to mean. As the resulting value is a dimensionless quantity, it can easily be compared with all evaluation criteria irrespective of any unit in which they were measured. Scores given by the evaluation committee were used when calculating the COV for technical proposals, interview, PPI, and RE. COV values for cost and schedule criteria were calculated using the cost proposals and schedule proposals submitted by each firm in a project.

4.3.4 KRUSKAL-WALLIS H TEST VARIABLES

The Kruskal-Wallis H test was used for testing three different data combinations. These combinations are as follows:

1) Independent variable: Six evaluation criteria for A/E projects

   Dependent variable: Median COV values

2) Independent variable: Six evaluation criteria per sub-factor for all project characteristics

   Dependent variable: Median COV values

3) Independent variable: Each evaluation criterion amongst every project characteristic's sub-factor

   Dependent variable: Median COV values

The first grouping of data consisted of the whole data sample without any classification, and the six criteria were compared against each other. The second grouping of samples compared the six evaluation criteria amongst each other for every sub-factor. For example, the six evaluation criteria were compared for pre-design projects to determine if there existed any statistical significant differentiation amongst each other. This was done for all sub-factors. The last grouping compared each evaluation criterion amongst sub-factors per project characteristics. For example, cost median COV was compared for pre-design and detailed design projects to determine if there existed a differentiation for the two sub-factors. Similarly, COV value for schedule, technical proposals, interview, PPI, and RE were compared for the two sub-factors. Then, the process was repeated for all project characteristics. All post-hoc analysis was conducted using pairwise comparison.

4.3.5 NORMALIZED COST AND SCHEDULE

Unlike other evaluation criteria, during the evaluation process, cost and schedule proposals were scored on the basis of lowest cost and lowest schedule. The lowest cost and the lowest schedule were awarded the highest scores while the rest of the cost and schedule proposals were scored using the inverse proportion method. All other evaluation criteria were scored on a scale of 0-100. In order to standardize cost and schedule proposals among projects of different size, a normalization technique was employed. Cost submissions were normalized in two ways: first, each competing consultant's cost was measured in terms of the percent difference from the lowest bid submitted for each project; second, each consultant's cost was measured as a percent difference from the average bid cost on the project. Each consultant's schedule proposal was normalized as the percent difference from the average schedule duration proposed per project. Normalization was necessary to compare cost and schedule proposals with other evaluation criteria, which otherwise would have been very difficult to compare due to inconsistent units (higher and lower volume projects and shorter and longer projects).

Cost as a percentage of lowest bid (% LB) for each project was calculated using: Cost (% LB) = (Lowest bid cost of the project– proposed cost of bidder) / Lowest bid cost of the project.

Similarly, cost as a percentage of average bid (% Avg.) for each project was calculated using: Cost (% Avg.) = (Average bid of the project- proposed cost of bidder)/ Average bid of the project. Last, schedule normalized (Norm.) for each project was calculated using: Schedule (Norm.) = (Least days for the project/ proposed days of the bidder). All these values were calculated in percentage.

4.4 EVALUATION PROCESS

4.4.1 TWO-ENVELOPE BV PROCUREMENT METHOD

All the projects within the dataset followed a virtually identical two-envelope BV evaluation process. The evaluation process started when the client released a Request for Proposal (RFP). Type of procurement and other evaluation criteria details were stated in the RFP. It also included templates for all the evaluation criteria, which were found in the attachment/appendix part of the RFP. An evaluation team, which generally was comprised of four to six members, evaluated submitted proposals individually and scored them on a scale of 1 to 10, with 1 denoting the lowest score and 10, the highest.

A pre-submittal conference was held for all expected bidders to familiarize them with the procurement process. The pre-submittal conference was generally held a week after the release of the RFP. A window of two weeks after the release was kept for a Question-Answer (Q/A) session. During this period, the owners' representative answered all questions and ambiguities put forward by the expected bidders. The deadline for submission of proposals was usually a week after the deadline for the Q/A session. No bids were accepted after the specified time. For BV procurement, bids were received in a two-envelope process. The first envelope consisted of all qualification documents (technical proposals, PPI, and RE) and the schedule proposal, while the second consisted of the bidders' cost proposal. Documents in the first envelope were then categorized into blind-evaluated and blind-not evaluated as shown in Figure 2 below.

Blind-evaluated documents consisted of all information submitted by the bidder, which were to be kept anonymous from the evaluation team for un-biased and transparent evaluations. Technical proposals were the only blind documents, while all other proposal documents were blind-not evaluated documents. The procurement officer checked all the blind documents for anonymity, name of the bidding company or its acronym, names of their past clients, affiliation with any institution which can help reveal their identity to the evaluator, mentions of their employee names, or any information which was not anonymous, and either deleted or highlighted the information in black before they were sent out to the evaluation team.

Technical proposals, which consisted of project capability, risk assessment plan, and value assessment plan, all followed a standard format in which the proposing bidders were asked to give a synopsis of the execution plan and enumerate all project specific risks and value added ideas. After anonymity checks from the procurement officer, a package that contained all blind documents and a score sheet to score the proposing bidder was sent to the evaluation team. These documents were the first things that were evaluated by the evaluation team. Members in the evaluation team were not revealed the identity of the bidder they were evaluating, for unbiased and transparent evaluation. The evaluation team were also told to evaluate the proposals independently so that they were not influenced by any other evaluator's decision or proclivity. The evaluators scored these plans on a scale of 1 to 10 and wrote down any comments they deemed necessary.

Meanwhile, the procurement officer inserted the scores for cost and blind-not evaluated proposals into the analysis file. Blind-not evaluated proposals were not evaluated by the evaluation team, but anonymity was still maintained in those documents. An analysis or evaluation file was an excel file with columns for each proposing bidder and their respective scores in each criterion.

A bidder-criterion matrix was formed when all the scores for each proposing bidder per evaluation criteria was keyed into the matrix. After the evaluation team was done evaluating the blind proposal documents, these were returned and were inserted in the analysis matrix file. At this point, the scores for all evaluation criteria were compiled, except the interviews. The procurement officer created a shortlist of the three to five highest scoring bidders to participate in the interview. Key personnel from each bidder's project team were interviewed by the owner's evaluation committee.

After the interview scores were inserted into the matrix, they were then transformed according to criteria weight. The summation of all the criteria weight assigned by the owner added up to 100%. The linear relationship model (LRM) was the method employed to change 1 to 10 scores into weighted scores. In this model, project cost and schedule were the only two evaluation criteria whose lowest value got the maximum weight, while all other evaluation criteria gave maximum weight for the highest scores in the evaluation process. All other scores were then weighted using the direct ratio method. The bidder with the highest weight was selected and awarded the contract.

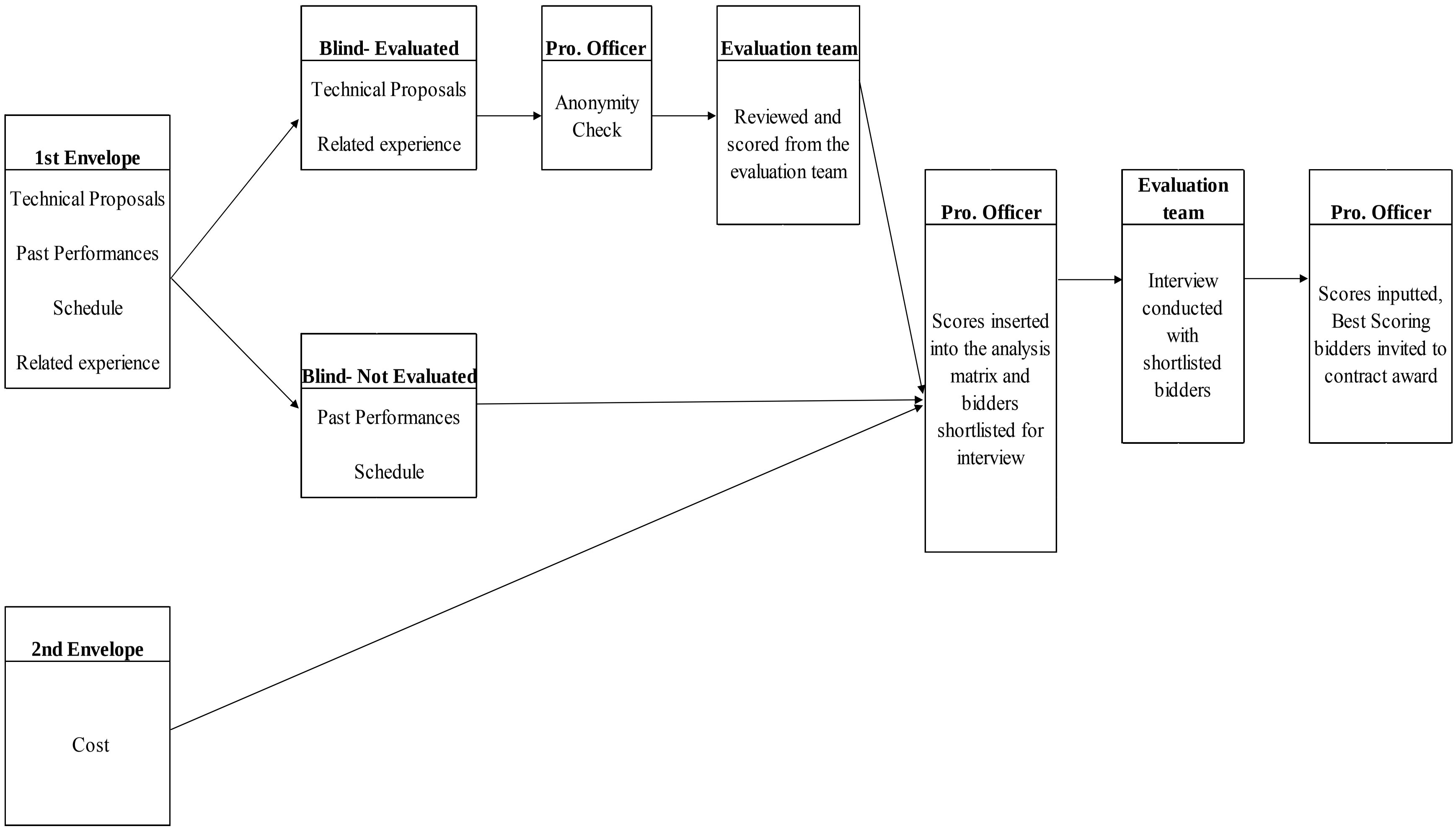

Figure 2. Evaluation Process for Two-envelope Best Value Procurement Method

4.4.2 QUALIFICATIONS-BASED PROCUREMENT METHOD

Qualifications-based procurement was similar to the two-envelope BV procurement method, except for the inclusion of cost factor. As no cost was considered when awarding contracts, QBS generally refers to RFP as RFQ, Request for Qualifications. It contains instructions to the bidder, evaluation criteria, evaluation methods, and standardized submittal attachments for evaluation criteria specified in the RFQ. A pre-submittal conference was held a week after the RFQ was released so that all prospected bidders could get acquainted with the procurement process.

After the RFQ was released, bidders submitted their proposals that were evaluated by the evaluation team. The evaluation process for QBS was a lot like the evaluation process for two-envelope BV procurement, except the process did not evaluate the cost criteria. Therefore, bidders were not required to submit a proposed cost within their submitted proposals. While the team was evaluating technical proposals and related experience, the procurement officer inserted schedule and past performance scores into the analysis/evaluation file. After the evaluation team evaluated all the bidders, these scores were sent to the procurement officer, who inserted all these scores and shortlisted the top three to five bidders (or more), depending on the company's policy or the owners' will, for the interview phase of the procurement.

All key personnel as mentioned in the RFQ were interviewed and scored by the evaluation team. Once all the shortlisted firms were interviewed and scored, these were then inserted by the procurement officer and the bidder with the highest weighted scores was invited to the negotiation stage. In the negotiation stage, both parties negotiated the cost of the project based on the scope of work. Therefore, it was very important for the owners to have a well-defined scope of work when negotiating the cost. If both parties agreed to the cost, the bidder was awarded the contract;

otherwise the owner could move to the next highest ranked bidder. Figure 3 shows the QBS procurement process.

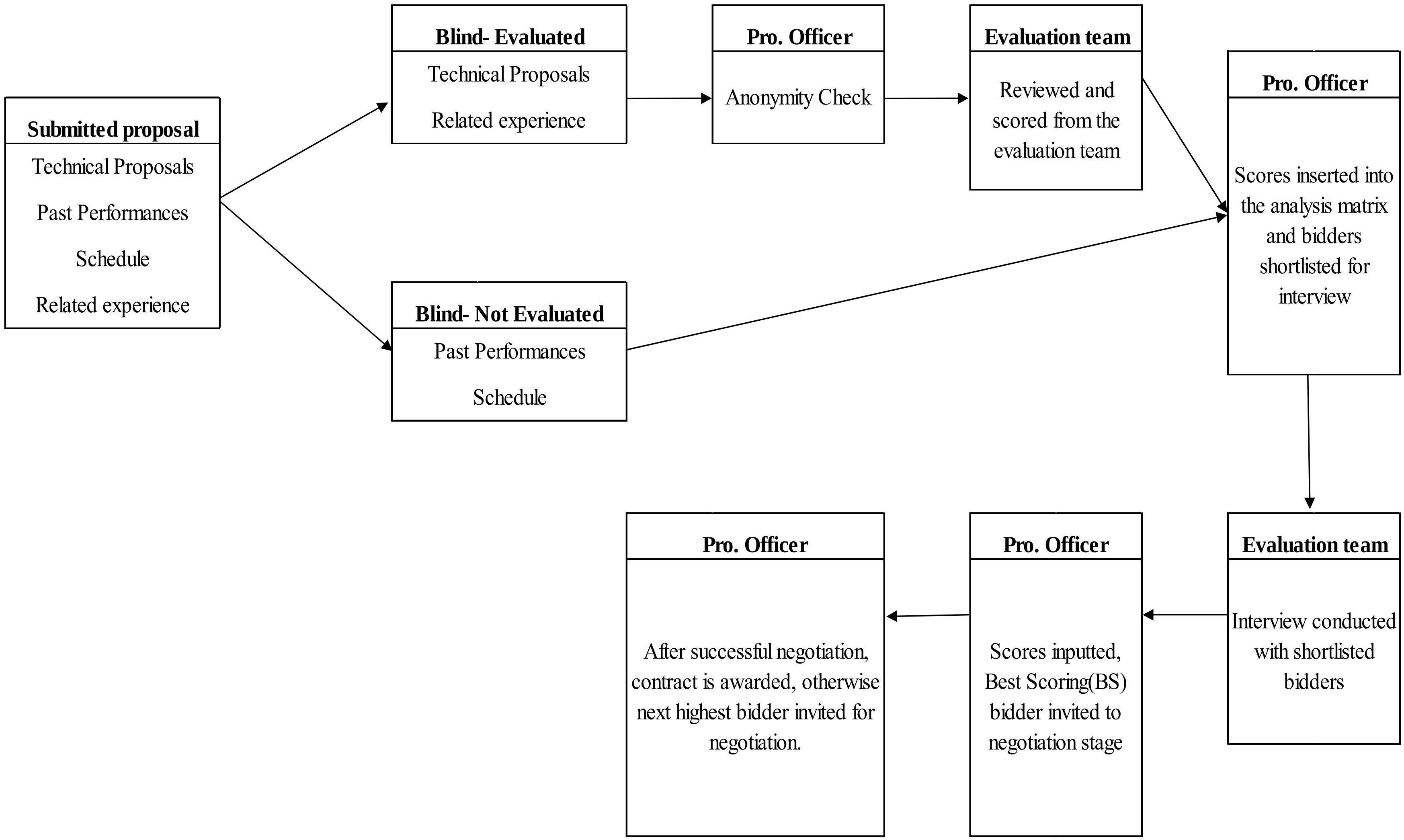

Figure 3. Evaluation Process for Qualifications-Based Procurement Method

4.5 DATA CHARACTERISTICS

Table 3 shows the distribution of weights for each evaluation criteria. Interviews and technical proposals were the top two highest weighted criteria with mean weights of 33.86% and 33.48%, respectively. High scores in one of the two evaluation criteria can greatly increase the chance of getting selected for the project. Projects where cost criterion was assigned 0% indicated that no cost criterion was used when selecting the bidder, hence these projects were procured using QBS. The minimum weight designated to cost criteria for the BV procurement method was 10% and went as high as 40%. The mean value for cost weight was 19.58%, which was well within the industry specification of using 30% when cost was used as an evaluation criteria when procuring A/E firms.

Table 3. Descriptive Analysis of Evaluation Criteria Weight

Criteria	N	Minimum	Maximum	Mean	Median	Std. Deviation
Cost	102	0*†	40.00%	19.58%	20.00%	6.80%
Schedule	82	5.00%	30.00%	13.96%	14.64%	5.56%
Technical Proposals	122	15.00%	55.00%	33.48%	35.00%	6.59%
PPI	121	5.00%	50.00%	19.90%	20.00%	9.14%
Interviews	54	20.00%	55.00%	34.86%	35.00%	7.82%
Related Experience	65	5.00%	26.67%	9.55%	6.67%	5.61%

*0 refers to projects procured via pure QBS
†For BV, the minimum weight was 10%

Moreover, schedule was assigned a mean weight of 13.69%, which was the second lowest mean weight amongst the six evaluation criteria. PPI criteria had a mean weight of 19.9%, with one anomaly in which a project assigned 50% to PPI criteria. RE was assigned the least weight in the evaluation process with a mean weight of 9.55%.

CHAPTER 5: METHOD OF ANALYSIS

Chapter 5 presents the methods of statistical analysis, both descriptive and inferential, used to analyze the data. Descriptive analysis, such as frequency analysis, rank matrices, and measure of central tendency, was used in order to answer the first research question of the thesis. The Kruskal-Wallis H test was used to answer the second, third, and forth research question. Spearman's correlation coefficient was used to answer the last research question. This chapter also provides justification for why these test were chosen to conduct analysis based upon the statistical assumptions required assumptions for the Kruskal-Wallis H test and Spearman's correlation coefficient. All these assumptions were met before conducting analysis.

5.1 DESCRIPTIVE ANALYSIS

One hundred and twenty-two A/E projects were analyzed, out of which 102 required cost proposals, eighty-two considered schedule proposals, 122 were technical proposals, 121 were past performances, fifty-four projects conducted interviews, and sixty-five considered RE as their evaluation criteria. To answer the first research question, frequency analysis was conducted for the characteristics of the selected bidders, including the frequency selections of the lowest bid (LB), best qualifications (BQ), and overall best score (BS). It consisted of frequency in percentage of selected lowest bidder (LB), best qualifications (BQ), and the best evaluation score (BS) for BV and QBS projects. Lowest bidder (LB) was denoted to bidders who had bid the lowest cost and were selected. Similarly, Best qualification (BQ) referred to bidders who were best in qualification and were selected. Lastly, Best evaluation score (BS) bidders were those who had the highest score for the evaluation. The BS was not necessarily the bidder with the lowest bid and the best qualifications, but there existed all four combinations. According to BV and QBS procurement, the BS was supposed to be awarded the contract, but this was not always the case. The data set encountered quite a few projects in which the owner chose either second or third ranked bidder when awarding the contract.

Furthermore, a Cost-Qualification (CQ) rank matrix was developed for BV procured projects, giving a detailed insight on the standings of the selected bidders. The matrix was a 4 x 4 square matrix with #1, #2, #3, and #4+ as its input variables for both cost and qualifications. A similar matrix was also developed for the six evaluation criteria. Fifteen such 4 x 4 matrices were developed, which showed the percentages for each rank and criterion combination. The matrix is provided as Table B.1 in Appendix B.

Descriptive analysis was performed on 122 projects to determine the selected bidders' ranking (SR), selected bidders' score (SS), and average project score (AS) for the six evaluation criteria. The average of these values was summarized in Table 6 with two more quality measures, "differential from the average bidder" ($\Delta_{avg.}$) and "differential from the lowest bidder" (Δ_{LB}). $\Delta_{avg.}$ and Δ_{LB} described the added value a selected bidder brought to the project when compared to the average bidder and lowest bidder. $\Delta_{avg.}$ and Δ_{LB} were calculated using the following formula: $\Delta_{avg.}$ = Average of (((SS-AS)/AS) x 100), and Δ_{LB} = Average of (((SS-LS)/LS) x 100), where LS stands for lowest bid score.

5.2 KRUSKAL-WALLIS H TEST

The Kruskal-Wallis H test, which is also known as the one-way ANOVA on ranks test, is a rank-based non-parametric test that was used to determine if there was a statistically significant difference between two or more groups of an independent variable (categorical). The independent variables were the six evaluation criteria, while COV was used as the dependent variable.

The main reason for selecting the Kruskal-Wallis H test over the one-way ANOVA test was that the Kruskal-Wallis H test was a non-parametric test. This means that the dependent variables do not have to be normally distributed. Furthermore, unlike the one-way ANOVA, the Kruskal-Wallis H test does not get affected by the outliers in the sample. Pairwise comparison was adopted as a post-hoc test, as it was the default post-hoc test for the Kruskal-Wallis H test in the Statistical Package for the Social Sciences (SPSS). Several assumptions must be met before conducting the Kruskal-Wallis H test, which are as follows:

1) Independent variables should be measured categorically with two or more groups.
2) Dependent variables to be continuous in nature.
3) Sample should have independence of observation.
4) Distribution of scores for each group of the independent variable was the same.

Every independent variable was categorical with either six evaluation criteria or two or more project characteristics' sub-factors. Coefficient of Variation (COV) was the dependent variable for the test and was continuous in nature. There was also no relationship between any of the independent variables, which satisfied the assumption for independence of observation. The last assumption was also satisfied when a box plot was plotted for all groups for independent variables. All the groups had the same shape of box plot and therefore, the last assumption was also met. All these box plots can be found in Appendix A, Figure A.1 to Figure A.23.

5.3 SPEARMAN'S CORRELATION COEFFICIENT

Spearman's correlation coefficient, also known as Spearman's rank-order correlation, calculated a coefficient r_s , pronounced as "rho," that measured the strength and direction of the association/relationship between two continuous or ordinal variables. The two main reasons for choosing Spearman's correlation coefficient are because it does not assume for the values to be normally distributed and does not get affected by the presence of outliers in the sample.

Nevertheless, there are still a few assumptions that must be met in order to perform Spearman's correlation. These are as follows:

1) Two variables to be continuous or ordinal.
2) Two variables represent a paired observation.
3) Monotonic relation exists between the two variables.

All the aforementioned assumptions were satisfied as technical proposals, interview, PPI, RE scores, normalized cost, and normalized schedule, which were continuous in nature and were paired. The last assumption assumed the variables to have a monotonic relation, which means that the variables had a linear relationship or not. This assumption was met by visual inspection of the scatter plot produced for the two variables. All these scatter plots can found in Appendix A, Figure A.24 to Figure A.30.

CHAPTER 6: RESULTS

Chapter 6 presents the results for descriptive analysis, the Kruskal-Wallis H test, and Spearman's correlation coefficient. The characteristics of the selected consultant and selection outcomes were described using the frequency analysis and rank matrices. In order to answer the second, third, and fourth research question, different evaluation criteria and project characteristics combination was used when conducting the Kruskal-Wallis H test and post hoc test. Lastly, this chapter present Spearman's correlation coefficient matrix which presents the relationship between the cost criterion to other qualifications criterion.

6.1 SELECTION OUTCOMES IN BEST VALUE PROCUREMENT OF A/E CONSULTANTS

Descriptive analysis was employed to determine the characteristics of the winning bidder. Table 4 shows the frequency with which the lowest bidder (LB), best qualifications (BQ), and best evaluation score (BS) were selected. When all BV procured projects were analyzed, it was found that 48.04% of the time, the lowest bidder emerged victorious, while 53.92% of the time, the best-qualified bidder was awarded the contract.

Table 4. Characteristics of Selected Bidders

Procurement method	N	Lowest Bid (LB)	Best Qualification (BQ)	Best Evaluation Score (BS)
BV	102	48.04%	53.92%	84.31%
QBS	20	n/a	100%	100%
Overall	122	40.16%	68.85%	88.52%

Moreover, the bidder with the best evaluation score, the top ranked bidder via the owners' procurement process, was not always awarded the project. The best evaluation score bidder was awarded the contract only 84.31% of the time. The remaining 15.69% of cases, the owners were incentivized to select the second ranked consultant (or third rank in just one case) based on the justification that the BS bidder was either too high or too low in cost.

With the inclusion of cost in the BV method, it was important to determine the occurrence of selected bidders that were both the lowest bid and the best qualification. Table 5 shows the arrangement of selected bidders by rank in cost and qualifications criteria. Cost criterion, as described earlier, was the cost proposal submitted by a bidder. Qualifications criteria was taken as a combination of technical proposals, interview, PPI, and RE. In a BV procurement process, the selected bidder had the best qualifications and also the lowest cost 21.57% of the time. Moreover, in 48.04% of the instances, the selected bidder was ranked first in cost or qualification and the top two in other. Lastly, 44.12% of the time, the lowest bidder was also ranked top three in qualifications; this explains why almost half the time in a BV procured project, the lowest bidder was selected.

Table 5. Cost-Qualification Matrix for Selected Bidders

	#1 Qualification	#2 Qualification	#3 Qualification	#4+ Qualification
#1 Cost	21.57%	17.65%	4.90%	3.92%
#2 Cost	8.82%	7.84%	0.98%	3.92%
#3 Cost	10.78%	0.00%	0.98%	1.96%
#4+ Cost	12.75%	2.94%	0.00%	0.98%

A detailed matrix for the six evaluation criteria was developed which is provided in Appendix 2, Table B.1. The matrix consists of fifteen combinations for the six evaluation criteria with their respective frequency (in percentage). The only notable result from the matrix was that 47.17% of the time, a firm ranked first in technical proposals was also ranked first in the interview stage of the procurement process.

Selected bidders' characteristics also include the standing of selected bidders in the six evaluation criteria and the score differential of the selected bidders from average bidder and lowest bidder. Table 6 shows that the selected bidder was placed, on average, first in interview and RE, second in cost, schedule, and technical proposals, third in past performance. Moreover, it was noted that the average scores of selected bidders in the PPI, interview, and RE were very high. PPI and related experience also had a very high average of average scores for the entire data.

Table 6. Differential between Selected Bidders and Competing Proposals

Criteria	Avg. Selected Bidders Ranking	Avg. Selected Bidders Score	Avg. of Average Scores	Differential from Average Bidder	Differential from lowest bidder
Cost	2.21	-	-	-15.74%	44.93%
Schedule	2.46	-	-	-15.79%	-5.11%
Technical Proposals	1.93	67.97	59.89	14.42%	34.25%
PPI	2.84	90.51	88.12	3.07%	15.15%
Interview	1.24	84.47	66.92	30.62%	180.07%*
Related Experience	1.20	90.66	90.54	0.52%	3.73%

*N= 17 projects in which lowest bid was shortlisted for the interview and was not the selected bidder

In the BV procurement system, the selected bidder always resulted in higher differentials when compared to the average bidder or the lowest bidder. Table 6 shows that in comparison with the average bidders' evaluation scores for the six evaluation criteria, the selected bidder was 15.74% less costly, had a 15.79% faster schedule, and 14.42%, 3.07%, 30.62%, and 0.52% more qualified in technical proposals, past performance, interviews, and related experience, respectively. Similarly, when compared to the lowest bidders' scores for the six evaluation criteria, the selected bidder was 44.93% more costly, had a 5.11% faster schedule, and 34.25%, 15.15%, 180.07%, and 3.73% more qualified in technical proposals, past performance, interviews, and related experience, respectively.

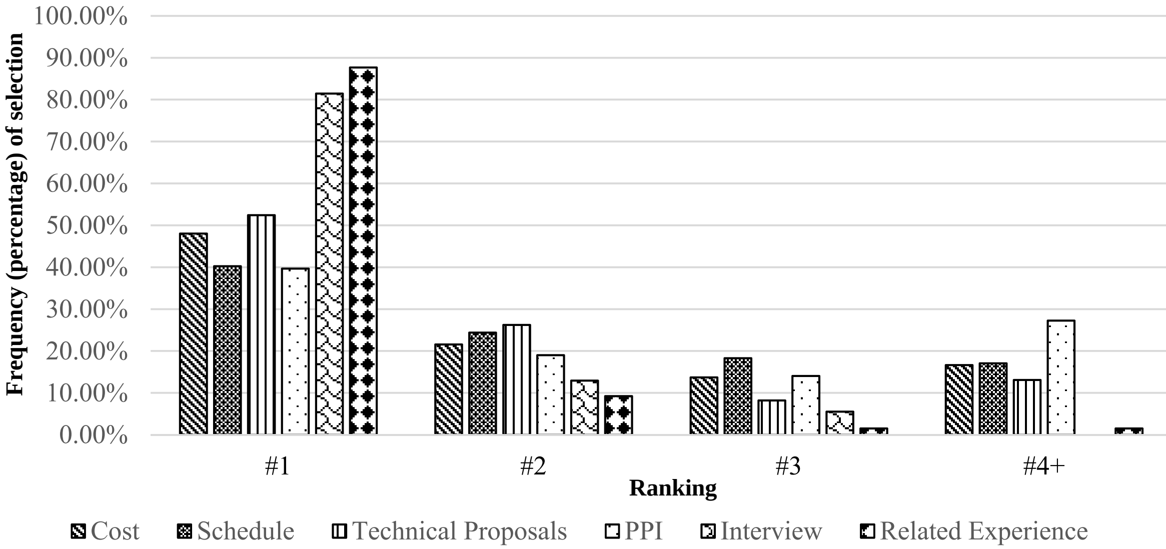

Figure 4. Ranking of Selected Bidders for All Evaluation Criteria

Figure 4 shows the rank frequency distribution of selected bidders per evaluation criteria in the six evaluation criteria. As shown in Figure 4, attaining first rank in the interview and RE criterion resulted in a very high probability of getting selected. Figure 4 also shows that 81% of the time, the selected bidder was ranked first in interviews compared to 13% and 6% for the second and third. No bidders were selected if they were ranked fourth or more. The firm ranked first in RE was selected 88% of the time, while the probability of selection decreased drastically if ranked lower. With 28%, PPI gives the highest percentage of a fourth or more ranked team being selected compared to all other evaluation criteria. Schedule and cost follow with 17% each.

6.2 EVALUATION CRITERIA EFFECTIVENESS IN DIFFERENTIATING CONSULTANT QUALIFICATIONS

6.2.1 KRUSKAL-WALLIS H TEST AMONGST ALL PROJECT CHARACTERISTIC

The Coefficient of Variation (COV) was used to measure of dispersion among competing consultant proposal elements within the data sample. This measure suited the need of the analysis to determine which evaluation criteria resulted in the greatest differentiation in evaluation scores between competing consultant proposals. Table 7 shows the descriptive statistics for the COV values for the six evaluation criteria. Schedule had the highest median COV value of 23.79%, followed by cost and interview with 22.50% and 17.10%, respectively. Technical proposals achieved a median COV value of 14.95%, while PPI achieved 4.80% and RE 0%.

Table 7. Descriptive Statistics for Coefficient of Variation

Criteria	N	Mean	Median	Std. Deviation	Minimum	Maximum	Calculated from
Cost	102	25.03%	22.50%	15.31%	0.00%	78.50%	$
Schedule	82	26.86%	23.79%	17.50%	0.00%	98.37%	Days
Technical proposals	122	17.14%	14.95%	11.56%	0.00%	68.60%	0-100
PPI	121	8.23%	4.80%	8.91%	0.00%	57.80%	0-100
Interview	54	24.13%	17.10%	20.76%	0.00%	90.73%	0-100
Related experience	65	7.62%	0.00%	15.87%	0.00%	70.70%	0-100

The first combination of the Kruskal-Wallis H test variable, where the independent variables were the six evaluation criteria and the dependent variable was the median COV values, were tested for statistical significance amongst the six evaluation criteria for competing A/E firms. The Kruskal-Wallis H test was used to determine if there were differences in median COV values between the six evaluation criteria: cost, schedule, technical proposals, PPI, interviews, and RE. Distribution of COV values was found to be similar for the six evaluation criteria, as assessed by visual inspection of boxplots, provided in Appendix A, Figure A.1. As shown in Table 8, median COV values for the six evaluation criteria were statistically significant between groups with $\chi 2$ (5) = 173.666 and p = 0.00. Post-hoc tests were performed to further investigate the statistical differences in COV between each of the evaluation criteria.

Table 8. KW-H Test for All Evaluation Criteria

Criteria	Sample Size (n)	Median COV	$\chi 2$-value	Degree of freedom (d.f)	*p*-value
Cost	102	22.50%	173.666	5	0.000**
Schedule	82	23.79%			
Technical proposals	122	14.95%			
PPI	121	4.80%			
Interview	54	17.10%			
Related Experience	65	0.00%			

**Statistically Significant at 0.01

Pairwise comparisons were performed using Dunn's (1964) procedure with a Bonferroni correction for multiple comparisons. This post hoc analysis, as shown in Table 9, revealed statistically significant differences in median COV values for p-values less than 0.05. As shown in the table, cost, schedule, and interviews with median COV values of 22.50%, 23.79%, and 17.10%, respectively, achieved the greatest differentiation in scores, while technical proposals with a median COV value of 14.95% gave a moderate differentiation. PPI and RE with median COV values of 4.80% and 0.00%, respectively, gave the least amount of differentiation in evaluation scores. This categorization was considered as the baseline classification for project characteristics test.

Table 9. Post-Hoc Testing for All Evaluation Criteria Combination

Factor 1	Factor 2	p-value
Related Experience	PPI	0.272
Related Experience	Technical Proposals	0.000**
Related Experience	Interview	0.000**
Related Experience	Cost	0.000**
Related Experience	Schedule	0.000**
PPI	Technical Proposals	0.000**
PPI	Interview	0.000**
PPI	Cost	0.000**
PPI	Schedule	0.000**
Technical Proposals	Interview	1.000
Technical Proposals	Cost	0.015*
Technical Proposals	Schedule	0.007**
Interview	Cost	1.000
Interview	Schedule	1.000
Cost	Schedule	1.000

*Statistically Significant at 0.05
**Statistically Significant at 0.01

Furthermore, the Kruskal-Wallis H test was also performed for the second combination of variables. The six evaluation criteria for every sub-factor in project characteristics, scope of work (pre-design and detailed design), project type (horizontal and vertical), procurement method (BV and QBS), prime vendor (architectural and engineering), project size (small, medium, and large), and project complexity (low, moderate, and high) were considered as independent variables and their respective median COV values as dependent variables. A box plot for each project characteristic's sub-factor is provided in Appendix A, Figure A.2 to Figure A.15. These tests were performed in order to determine if there existed a differentiation in the six evaluation criteria for each project characteristic's sub-factor, and if so, how different the classification was for each sub-factor compared to the baseline classification.

Every sub-factor was found to be statistically significant at p-value 0.05. When post-hoc (pairwise comparison) tests were conducted for each of these sub-factors, a table similar to Table 9 was generated showing which two evaluation criteria were statistically significant (these tables are provided in Appendix B, Table B.2 to Table B.15.) Based on the results of pairwise comparison of the six evaluation criteria for each sub-factor, these were then classified into greatest, moderate, and least differentiation. As shown in Table 10, except for a few deviations, no major changes were noticed in the results, and the sub-factors' categorization was found to be consistent with the baseline classification.

Table 10. KW-H Test and Classification for Project Characteristics

Project Characteristic	Sub-factors	Degrees of Freedom (d.f)	χ2-value	*p*-value	Greatest Differentiation	Moderate Differentiation	Least Differentiation
All selection criteria (Baseline Classification)					C, S, I	TP	PPI, RE
Scope of work	Pre-Design	5	28.294	0.000*	C, S, TP	I	PPI, RE
	Detailed Design	5	148.308	0.000*	C, S, I	TP	PPI, RE
Project type	Horizontal	5	40.718	0.000*	C, S, TP, I	-	PPI, RE
	Vertical	5	142.200	0.000*	C, S, I	TP	PPI, RE
Procurement method	QBS	3	24.771	0.000*	S, I, TP	PPI	-
	BV	5	164.097	0.000*	C, S, I	TP	PPI, RE
Prime vendor	Architectural	5	117.782	0.000*	C, S, I	TP	PPI, RE
	Engineering	5	74.779	0.000*	C, S, TP, I	-	PPI, RE
Size of project	Small	4	88.224	0.000*	C, S	TP, PPI	RE
	Medium	5	58.174	0.000*	C, S, TP, I	-	PPI, RE
	Large	5	45.105	0.000*	C, S, TP, I	PPI	RE
Project complexity	Low	5	53.198	0.000*	C, S, TP, I	-	PPI, RE
	Moderate	5	66.986	0.000*	C, S, TP	I	PPI, RE
	High	5	47.728	0.000*	C, S	TP, I	PPI, RE

C=Cost, S=Schedule, TP=Technical Proposals, I=Interview, PPI=Past Perf. Inf., RE=Related Exp.
*Statistically Significant at 0.05

6.2.2 KRUSKAL-WALLIS H TEST AMONGST INDIVIDUAL PROJECT CHARACTERISTIC

The Kruskal-Wallis H test was also conducted individually for each project characteristic. This time, every sub-factor in a particular project characteristic was compared in order to determine any differentiation for an evaluation criterion. This was done for the six evaluation criteria.

A Kruskal-Wallis H test was run to determine if there were differences in median COV values for cost, schedule, technical proposals, PPI, interview, and RE between the scope of work's sub-factors, pre-design and detailed-design. The distribution of median COV values was found to be similar for both the sub-factors, as assessed by visual inspection of a boxplot found in Appendix A, Figure A.16. As shown in Table 11, the median interview COV values for pre-design projects (10.80%) and detailed design projects (17.68%) were found statistically significant with $\chi 2$ (1) = 3.847 and p = 0.05. This shows that interview scores for detailed design projects achieved a higher differentiation when compared to pre-design projects. All other evaluation criteria were either not significant or were too limited in sample size to conduct any analysis.

Table 11. KW-H Test Evaluation Criteria Differentiation Based on Scope of Work

Criteria	Pre-Design		Detailed Design		Degrees of Freedom (d.f)	$\chi 2$-value	*p*-value
	N	Median COV	N	Median COV			
Cost	12	25.05%	90	22.50%	1	0.770	0.380
Schedule	9	26.61%	73	23.63%	1	0.143	0.705
Technical proposals	19	14.80%	103	15.10%	1	0.001	0.980
PPI	19	3.60%	102	5.60%	1	0.001	0.974
Interview	9	10.80%	45	17.68%	1	3.847	0.05*
Related Experience	12	0.00%	53	0.00%	1	0.400	0.527

*Statistically Significant at 0.05

Another Kruskal-Wallis H test was run to determine if there were differences in median COV values for cost, schedule, technical proposals, PPI, interview, and RE between project types' sub-factors: horizontal and vertical projects. The distribution of median COV values was found to be similar for both the sub-factors, as assessed by visual inspection of a boxplot found in Appendix A, Figure A.17. As shown in Table 12, the median schedule COV values for horizontal projects (48.85%) and vertical projects (23.13%) were found statistically significant with $\chi 2(1) = 9.554$ and $p = 0.002$. This shows that schedule scores for horizontal projects achieved higher differentiation than vertical projects.

Table 12. KW-H Test Evaluation Criteria Differentiation Based on Project Type

Criteria	Horizontal		Vertical		Degrees of Freedom (d.f)	χ2-value	*p*-value
	N	Median COV	N	Median COV			
Cost	12	15.20%	90	23.75%	1	2.991	0.084
Schedule	5	48.85%	77	23.13%	1	9.554	0.002**
Technical proposals	25	16.30%	97	14.80%	1	1.011	0.742
PPI	24	3.55%	97	6.30%	1	0.109	0.171
Interview	19	14.65%	35	17.68%	1	1.873	0.315
Related Experience	4	0.00%	61	0.00%	1	0.030	0.863

**Statistically Significant at 0.01

A Kruskal-Wallis H test was run to determine if there were differences in median COV values for cost, schedule, technical proposals, PPI, interview, and RE between procurement methods' sub-factors, QBS and BV. The distribution of median COV values was found to be similar for both the sub-factors, as assessed by visual inspection of a boxplot found in Appendix A, Figure A.18. As shown in Table 13, the median interview COV values for QBS projects (10.89%) and BV projects (24.21%) were found statistically significant with $\chi 2\ (1) = 13.454$ and $p = 0.000$. Therefore, interview scores for BV procured projects gave a higher differentiation than QBS procured projects. All other evaluation criteria were either not significant or were too limited in sample size to conduct any analysis.

Table 13. KW-H Test Evaluation Criteria Differentiation Based on Procurement Type

Criteria	QBS		BV		Degrees of Freedom (d.f)	χ2-value	*p*-value
	N	Median COV	N	Median COV			
Cost	-	-	102	22.50%	-	-	.
Schedule	2	39.15%	80	23.50%	1	2.170	0.141
Technical proposals	20	16.85%	102	14.70%	1	0.175	0.676
PPI	19	4.80%	102	4.35%	1	0.040	0.842
Interview	19	10.89%	35	24.21%	1	13.454	0.000**

*Statistically Significant at 0.05
**Statistically Significant at 0.01

A Kruskal-Wallis H test was run to determine if there were differences in median COV values for cost, schedule, technical proposals, PPI, interview, and RE between prime vendors' sub-factors: architectural and engineering. The distribution of median COV values was found to be similar for both sub-factors, as assessed by visual inspection of a boxplot found in Appendix A, Figure A.19**.** As shown in Table 14, the median cost COV values for architectural projects (20.65%) and engineering projects (29.95%) were found statistically significantly different with $\chi 2$ (1) = 5.945 and p = 0.015. This shows that cost proposals for engineering projects achieved a higher differentiation of scores when compared to architectural projects. All other evaluation criteria were either not significant or were too limited in sample size to conduct any analysis.

Table 14. KW-H Test Evaluation Criteria Differentiation Based on Prime Vendor

Criteria	Architectural		Engineering		Degrees of Freedom (d.f)	χ2-value	*p*-value
	N	Median COV	N	Median COV			
Cost	72	20.65%	30	29.95%	1	5.945	0.015*
Schedule	64	23.50%	18	27.04%	1	0.813	0.367
Technical proposals	77	14.00%	45	17.40%	1	3.002	0.083
PPI	77	6.10%	44	4.75%	1	0.024	0.876
Interview	27	18.27%	27	14.61%	1	0.022	0.883
Related Experience	47	0.00%	18	0.00%	1	0.042	0.838

*Statistically Significant at 0.05

A Kruskal-Wallis H test was run to determine if there were differences in median COV values for cost, schedule, technical proposals, PPI, interview, and RE between the size of projects' sub-factors: small, medium, and large. The distribution of median COV values was found to be similar for the three sub-factors, as assessed by visual inspection of a boxplot found in Appendix

A, Figure A.20**.** As shown in Table 15, the median COV values were statistically significantly different between groups for technical proposals, $\chi 2$ (1) = 2.255, p = 0.022, and RE, $\chi 2$ (2) =19.187, p = 0.000. All other evaluation criteria were either not significant or were too limited in sample size to conduct any analysis. Moreover, a post-hoc test (pairwise comparison) was performed in order to determine where the significance in the three sub-factors lay.

Table 15. KW-H Test for Evaluation Criteria Differentiation Based on Size of Project

Criteria	Small		Medium		Large		Degrees of Freedom (d.f)	$\chi 2$-value	*p*-value
	N	Median COV	N	Median COV	N	Median COV			
Cost	37	27.90%	30	22.50%	35	18.20%	2	3.604	0.165
Schedule	35	23.63%	29	23.08%	16	28.48%	2	6.028	0.324
Technical proposals	37	12.20%	30	14.30%	35	16.90%	2	2.255	0.022*
PPI	37	2.50%	30	2.50%	35	6.60%	2	7.644	0.57
Related Experience	30	0.00%	18	0.00%	16	10.00%	2	19.187	0.000**

*Statistically Significant at 0.05
**Statistically Significant at 0.01

The results of the post-hoc test, as shown in Table 16 below, revealed a statistically significant difference in median COV values for technical proposals between small projects (12.20%) and large projects (16.9%) at p = 0.022. Technical proposal scores achieved a higher differentiation for large projects compared to small. A statistically significant result was also found for median COV values for RE between small projects (0.00%) and medium projects (0.00%), with large projects (10.0%) at p = 0.002 and 0.00, respectively. In both the cases, RE scores achieved greater differentiation for large projects compared to small and medium.

Table 16. Post-Hoc Test for Size of Project

Criteria	Factor 1	Factor 2	*p*-value
	Small	Medium	-
Cost	Medium	Large	-
	Small	Large	-
	Small	Medium	-
Schedule	Medium	Large	-
	Small	Large	-
	Small	Medium	0.817
Technical Proposals	Medium	Large	0.379
	Small	Large	0.018*
	Small	Medium	-
PPI	Medium	Large	-
	Small	Large	-

	Small	Medium	1.000
Interview	Medium	Large	0.060
	Small	Large	1.000
	Small	Medium	1.000
Related Experience	Medium	Large	0.000**
	Small	Large	0.002**

*Statistically Significant at 0.05
**Statistically Significant at 0.01

A Kruskal-Wallis H test was run to determine if there were differences in median COV values for cost, schedule, technical proposals, PPI, interview, and RE between project complexity's sub-factors: low, moderate, and high complex projects. The distribution of median COV values was found to be similar for the three sub-factors, as assessed by visual inspection of a boxplot found in Appendix A, Figure A.21. As shown in Table 17, the median COV values were statistically significantly different between groups for schedule, $\chi2$ (2) = 8.177 at p = 0.017. All other evaluation criteria were either not significant or were too limited in sample size to conduct any analysis. Moreover, a post-hoc test (pairwise comparison) was performed in order to determine where the significance in the three sub-factors lay.

Table 17. KW-H Test between All Criteria and Project Complexity

Criteria	Low		Moderate		High		Degrees of Freedom (d.f)	$\chi2$-value	*p*-value
	N	Median COV	N	Median COV	N	Median COV			
Cost	25	17.20%	30	23.90%	47	24.70%	2	2.441	0.295
Schedule	18	35.67%	25	28.37%	39	20.94%	2	8.177	0.017*
Technical proposals	35	14.00%	31	14.80%	56	16.60%	2	0.096	0.953
PPI	34	4.10%	31	6.30%	56	5.60%	2	0.738	0.691
Interview	19	16.95%	11	16.67%	24	17.98%	2	0.054	0.973
Related Experience	15	0.00%	20	0.00%	30	0.00%	2	0.145	0.930

*Statistically Significant at 0.05

The results from the post-hoc test, as shown in Table 18 below, revealed a statistically significant difference in median COV values for schedule between high complexity projects (20.94%) and low complexity projects (35.67%) at p = 0.015. This means that schedule proposals attained higher differentiation for low complex projects compared to high complex projects.

Table 18. Post-Hoc Test for Project Complexity

Criteria	Factor 1	Factor 2	*p*-value
	Low	Moderate	-
Cost	Moderate	High	-
	High	Low	-
	Low	Moderate	0.507
Schedule	Moderate	High	0.426
	High	Low	0.015*
	Low	Moderate	-
Technical Proposals	Moderate	High	-
	High	Low	-
	Low	Moderate	-
PPI	Moderate	High	-
	High	Low	-
	Low	Moderate	-
Interview	Moderate	High	-
	High	Low	-
	Low	Moderate	-
Related Experience	Moderate	High	-
	High	Low	-

*Statistically Significant at 0.05

6.2.3 OWNER PROVIDED BUDGET AND SCHEDULE

After a Kruskal-Wallis H test was conducted for the level of differentiation achieved by different evaluation criteria, results showed that the cost and schedule proposals both achieved very high differentiation of COV values. A follow up analysis was conducted with two new hypotheses that helped in explaining the earlier results. The data was divided into two more classifications, 'whether the owner provided a design budget or not' and 'whether the owner provided a schedule for the project or not.'

A Kruskal-Wallis H test was run to determine if there were differences in median COV values for 'yes' or 'no' for the owner-provided budget or schedule. The distribution of median COV values was found to be similar for both the groups, as assessed by visual inspection of a boxplot found in Appendix A, Figure A.22 and A.23. As shown in Table 19, the median COV values were statistically significant for both the evaluation criteria: cost, $\chi 2$ (1) = 19.07, p = 0.00, and schedule, $\chi 2$ (1) = 4.89, p = 0.027. This means that when the owner provided some kind of cost and schedule information, the cost and schedule proposals had lower COV values compared to when the owner provided no information at all regarding the cost and schedule.

Table 19. KW-H Test for Owner-Provided Budget and Schedule against Median Cost and Schedule COV

Evaluation criteria	Owner Provided Project Budget and Schedule				Degrees of Freedom (d.f)	$\chi 2$-value	*p*-value
	Yes		No				
	N	Median COV	N	Median COV			
Cost	78	20.30%	24	37.45%	1	19.07	0.000*
Schedule	71	25.63%	11	34.82%	1	4.89	0.027*

*Statistically Significant at 0.05

6.3 RELATIONSHIP BETWEEN CONSULTANT QUALIFICATIONS AND BID COST

Cost was found to have no practically significant correlation any qualifications-based evaluation criteria. Spearman's correlation coefficient was used in order to check if a relationship existed between cost proposal and qualifications criteria. Monotonic assumption between the variables was satisfied. Monotonic assumption was checked by the visual inspection of the scatter plot, which can be found in Appendix A, Figure A.24 to Figure A.30. Table 20 shows the Spearman's correlation coefficient for each combination of evaluation criteria. Cost (% Avg.) showed a slight negative correlation with schedule (Norm.) with Spearman's correlation (rho) r_s = -0.177, which was significant at 0.01. Similarly, cost (% LB) was also found to be statistically significant to schedule (Norm.) and PPI with Spearman's correlation (rho) r_s = -0.287 and -0.117, respectively, both significant at 0.01. Schedule (Norm) was also significant to PPI with Spearman's correlation (rho) r_s = 0.113, which was significant at 0.05. Technical proposals were found to be significant to PPI and related experience with Spearman's correlation (rho) r_s = 0.095 and 0.102, respectively, both significant at 0.05, while technical proposals were significant to interview with a moderately positive Spearman's correlation (rho) r_s = 0.426, significant at 0.01. The extremely weak correlation coefficients indicate the associations are of no practical impact for industry practitioners. No other qualifications criteria showed significant result with cost.

Table 20. Spearman's Correlation Coefficient for All Evaluation Criteria Combination

		1	2	3	4	5	6	7
1	Cost (% Avg.)	1.000						
2	Cost (% LB)	-	1.000					
3	Schedule (Norm.)	-0.177**	-0.287**	1.000				
4	Technical Proposals	0.061	-0.015	0.088	1.000			
5	PPI	0.029	-0.117**	0.113*	0.095*	1.000		
6	Interview	0.075	0.165	0.006	0.426**	0.056	1.000	
7	Related Experience	-0.002	0.026	0.079	0.102*	-0.180	0.296	1.000

*Correlation is significant at the 0.05 level (2-tailed)

**Correlation is significant at the 0.01 level (2-tailed)

CHAPTER 7: DISCUSSION, CONCLUSIONS, AND CONTRIBUTIONS

The last chapter of the research present the discussion to the result of test conducted in the previous chapter. The characteristic of the selected bidder, differentiation achieved by the six evaluation criteria individually and for different project characteristics, and lastly the cost and qualification relationship are discussed in detail. Conclusions were drawn based on the results and the thesis contribution to the body of knowledge and industry practitioners were also described in this chapter. At the end of this chapter limitation to the thesis and future research was also discussed.

7.1 SELECTION OUTCOMES IN BEST VALUE PROCUREMENT OF A/E CONSULTANTS DISCUSSIONS

For BV procured projects, almost half the time, the BQ consultant was selected, compared to qualifications- based procurement system. This indicated that owners today generally want better-qualified consultants who can understand the scope of work, refine it, look for every possible risk involved, and add as much value to the project as possible. Although the LB firms were awarded the contract almost 48% of the time, one would think that even with low weight assigned to cost criterion, the lowest bidder was still able to get selected half the time. At the same time, it should also be noted that out of the 48%, almost 21.5% of the time the lowest bidder was also the best in qualification, therefore, only 26.5% of the bidders were the lowest bidder and not the best in qualifications. However, in that 26.5%, around 7% of the time, the owner selected the lowest bidder because the BS bidder was too high in cost, while the lowest bidder was second (or third in one project) in rank in qualifications. To be precise, only 19.5% of the time, selected bidders were actually the LB and were the BS in the procurement process. This was just below the mean cost weight of 20%. Figure 5 shows a Venn diagram demonstrating the distribution of projects (in percentages) for LB and BQ.

Moreover, with mean weight assigned to qualifications criteria to be almost 74%, the BQ was awarded the contract only 53.92% of the time. The reasons for not selecting the BQ can again be linked back to the owners' selection by will. The owners did not select the BQ (which was also the BS) 11% of the time, and turned to the bidder second or third in rank, due to a very high or low cost proposal from the BS bidder. The cost criterion still has an influence on the owners as they are inclined to pick the lowest bidder when they are second or third in place in total evaluation scores. This leads to the conclusion that inclusion of cost as an evaluation criterion does have a disproportionate effect on selection outcomes, such that the owners do trend towards the selection of lowest bid or non-highest qualifications.

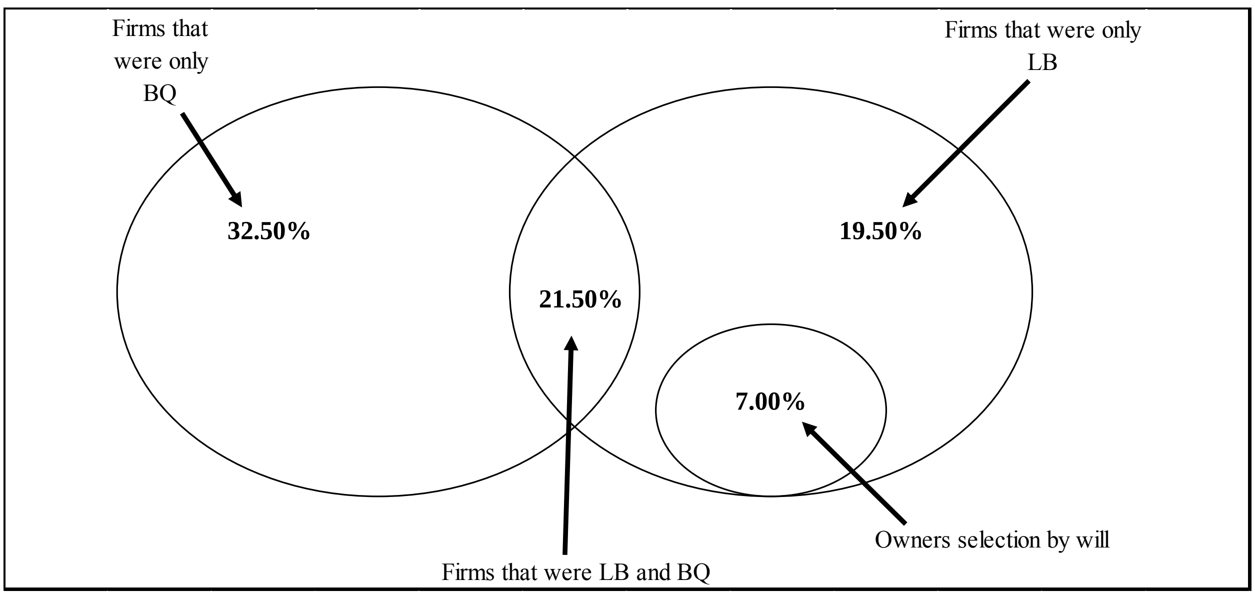

Figure 5. Venn diagram for Selected BQ and LB Firms

At first glance, consultants may fear that the owners were diverting their selection in favor of lower-cost options, yet closer inspection reveals the owners' behavior to be rational. For example, of the cases in which the best evaluation score bidder was not selected, three-quarters of the time the bidder had significantly high cost proposals. On average they were 27.6% higher costs than the second rated bidder, with a 4.63% differential in total evaluation scores, essentially indicating that the qualifications between the top two firms were identical (or even favored the second-ranked firm). The remaining 25% of bidders had significantly low cost proposals, on average 42.5% less than the second rated bidder, with only a 4.97% differential in total evaluation score. The owners in those cases chose to select the higher cost proposals that they felt were more realistic. Out of the twelve cases where the owner chose to select a bidder other than best evaluation score due to the cost proposal being very high, only one case had an exception in which the owner selected the third in rank instead of the second.

Adding more to the characteristics of the selected bidder, when cost and qualification matrices were developed for the six evaluation criteria, as shown in Appendix B, Table B.1. The results indicated a high percentage of selection of A/E firms when ranked number one in interview and technical proposals. This relationship was further bolstered by the result of Spearman's correlation, which showed a moderately positive correlation between the interview and technical proposals scores. Other significant results include a high selection percentage of A/E firms when ranked first in RE and also ranked first in any other evaluation criteria, generally greater than 35%. With many bidders getting the highest score in RE for a project, the selected bidder was most likely to be ranked one with a shared high score. Furthermore, Figure 4 supports this claim as it shows that 88% of the time the selected bidders were ranked first in RE.

As shown in Table 6, selected bidders had high scores for interview, PPI, and RE. High scores in interview can correlate to the most qualified bidders being shortlisted for the interview stage. Due to tough competition, selected bidders have to have a very high score so that they can differentiate themselves from other bidders. Interviews were therefore considered a crucial part of the procurement process. The reason for the selected bidders achieving high scores in PPI criterion was because PPI surveys were completed and scored by the bidders' past clients. Generally, bidders tend to submit surveys that have scored them the most (mostly in a range of 90 to 100).

Figure 4 also shows that firms first in interview and RE had a very high percentage of selection. In the interview process, only the highest qualified firms were selected, and on average, 35% weight was assigned to the interview stage. Therefore, firms ranked first in interview received the highest weight, which resulted in becoming the highest ranked bidder.

As discussed earlier, for RE, almost all the bidders achieved the highest score, which resulted in every bidder being the top ranked firm for RE in that project. Hence, it was very likely that firms selected were also ranked first in RE criterion. Lastly, it was noticed that the firms ranked fourth or more in PPI were still awarded the contract. The reason for this anomaly was that the average scores of past performances were high. These scores were generally clustered in a very short range (90 to 100); therefore, ranking in PPI does not help differentiate between a high scoring bidder and a low scoring bidder.

7.2 EVALUATION CRITERIA EFFECTIVENESS IN DIFFERENTIATING CONSULTANT QUALIFICATIONS

7.2.1 KRUSKAL-WALLIS H TEST AMONGST ALL PROJECT CHARACTERISTICS

As shown in Table 8, the cost proposal, schedule proposal, and interview evaluation criteria gave the most differentiation based on their respective median COV values. It was not always that the owners provided budget or schedule information for the project, therefore, it was entirely up to the discretion of the bidders to propose their own estimated cost and schedule. Since every bidder proposed bids based on their own design, a higher differentiation existed for cost proposal and schedule proposal amongst all the proposing bidders. Furthermore, interviews were generally designed in a manner where a detailed review of the submitted technical proposals were discussed.

Therefore, the evaluation team scored bidders based on the specifics of the project, resulting in high scores for highly qualified firms and vice versa. Technical proposals gave a moderate differentiation for the COV values, while PPI and RE achieved the least differentiation. As PPI was scored by past clients for bidders' performance in their project, bidders only submitted surveys which scored them highly, resulting in less differentiation. Moreover, the majority of the bidders in a project got the highest RE score, which gave an almost zero standard deviation, resulting in zero or very low COV values. The median COV value for RE was zero, which indicated that almost half the projects in the data or even more had zero as their COV value.

A Kruskal-Wallis H test with a post-hoc (pairwise comparison) test was conducted for different project characteristics' sub-factors. The results showed that COV values for cost proposal, schedule proposal, and interviews gave the greatest differentiation, while technical proposals achieved a moderate differentiation, and PPI and RE gave the least differentiation. All sub-factors for every project characteristic, with a few deviations, were found to be consistent with the baseline classification. Therefore, $H1_1$ was accepted; that different evaluation criteria will result in various levels of differentiation between competing A/E firms.

7.2.2 KRUSKAL-WALLIS H TEST AMONGST INDIVIDUAL PROJECT CHARACTERISTICS

As shown in Table 11, interview scores for detailed design projects gave a higher differentiation when compared to pre-design projects. This was because in detailed design projects, the scope of work was very well defined, making the interview process very accurate and precise. On the contrary, pre-design projects' scope of work was not clearly defined as the project was still in the conceptual stages. This led to the interview stage being very general, which provided the interviewee with an opportunity to provide marketing information rather than talk about the specifics of the scope of work, resulting in a very low differentiation in interview scores.

Schedule proposal criterion for horizontal projects gave a higher differentiation of scores when compared to vertical projects, as shown in Table 12. Even though the sample only contained procurement of AE firms, the significance in schedule was very hard to justify. One possible reason for this anomaly can be the smaller sample size of horizontal projects, which was five, compared to vertical projects, which was seventy-seven.

For interview, BV procured projects achieved a higher differentiation of scores compared to QBS procured projects. For QBS projects, firms who were interviewed were shortlisted purely on the basis of qualifications. Therefore, it was highly probable that all the firms shortlisted were highly qualified and therefore the range in scores was less. While for BV procured projects, because cost was also considered when shortlisting for the interview stage, firms that were not very qualified but bid low cost could still be shortlisted, which resulted in a higher differentiation of interview scores.

Projects with engineering firms as the prime vendor gave a higher differentiation of cost scores compared to projects with architectural firms as the prime vendor. The reason for such a significance was hard to explain, as both entities were design professionals and both were compensated based on their experience and services they rendered.

For the size of project, a Kruskal-Wallis H test with pairwise comparison as the post-hoc test yielded significant results for technical proposals and related experience. Technical proposal scores for small projects showed a lesser differentiation when compared to large projects. Small projects tend to have less complexity when compared to large projects, hence the risk plan, execution plan, and value added plan were very generic. Therefore, it was easier for the evaluation team to score all bidders in a closer range. On the other hand, technical proposals for large projects were more convoluted and therefore, the proposals submitted vary in risk plan, execution plan, and value added plan amongst every bidder. This resulted in a higher differentiation in technical proposal scores. Moreover, large projects showed a higher differentiation in RE scores when compared to small and medium projects. The results are not considered a good measure, as RE's median COV values for small and medium projects were both 0.00%.

Except schedule, no other evaluation criterion showed a significant result for project complexity. Further analysis using pairwise comparison showed that low complex projects achieved a greater differentiation in schedule scores compared to high complex projects. These results were surprising and counterintuitive because generally, low complex projects tend to have a straightforward project schedule compared to high complex projects, which are mostly multi-year projects, hence the schedule bids are assumed to have greater variations. One reason could be the shorter duration of low complex projects compared to high complex projects. As low complex projects have a comparatively shorter duration, a project with a smaller range (highest schedule day – lowest schedule day) can lead to a very high value of COV. Whereas, the same range for high complex projects, which have longer durations (typically in years), could lead to lower values of COV. Lastly, no statistically significant result was found for moderate complex projects from either of the other two sub-factors.

Although not every evaluation criteria was found to achieve a differentiation in project characteristics, each of the project characteristics showed a significant result in at least one of the evaluation criteria. Therefore, $H2_1$ was accepted; that the evaluation criteria will achieve various levels of differentiation based upon separate project characteristics.

7.2.3 OWNER PROVIDED BUDGET AND SCHEDULE

As described in earlier sections, owner-provided information regarding the project budget and schedule days was found to be significant for median COV values for cost proposals and schedule proposals. When the owner provided some form of budget information, be it design budget, construction budget, or both, cost proposals came in with lesser variations compared to projects in which the owner did not provide any information. Bidders tend to submit cost proposals in close proximity to the provided budget of the owner in order to increase their chances of getting selected. When the owner did not provide any information regarding the budget, bidders quoted a cost based on their own design and schedule. This led to higher differentiation in cost proposal scores.

Moreover, when the owner provided any information regarding the project schedule, be it as detailed as anticipated design start date, bid date, and construction start date, or as simple as number of days/years for the project, schedule proposals came in with lesser variations compared to when the owner provided no information at all. The reason for differentiation was the same as when the owner did not provide any cost information. The bidders were incentivized to bid the schedule days in close proximity to the owners' provided schedule. On the contrary, when the owner provided no information for the project schedule, it was up to the bidders to provide a schedule, for as many or as few days as they wanted, based on their design. Therefore, $H3_1$ was accepted, that the project budget provided will result in lesser variation of cost proposal (lower values of cost Coefficient of Variation (COV)). $H4_1$ was also accepted, that the provided schedule will result in lesser variation of schedule proposals (lower values of schedule days Coefficient of Variation (COV)).

7.3 RELATIONSHIP BETWEEN CONSULTANT QUALIFICATIONS AND BID COST

With all Spearman's correlation (rho) values shown in Table 20, only a few criteria were found to have a relation with other criteria. Technical proposals and interview scores showed a moderately positive Spearman's correlation value of 0.426, which was significant at p-value 0.01, while all other correlation coefficients were either less than 0.3 or greater than -0.3. These correlations were practically insignificant because of such a weak correlation coefficient. Meanwhile, the relationship between technical proposals and interview showed a moderate correlation because when technical proposals were accurate and well written, evaluation teams scored them very high. Also, interviews were designed to interact with individuals who had technical relevance with the project. These individuals were also the ones responsible to write or assist in developing the technical proposals for their company. During interviews, key individuals discussed the risks associated, value assessment, and execution plan of the project in greater detail, which led to high scores in the interview for well-written technical proposals. Traditionally, marketing teams were sent to the interview stage, highlighting only the market aspects of the company rather than talking about the technicalities of the project, which made it hard for the owners to select the most qualified firm based on project specific technicalities.

Cost was compared to other evaluation criteria using two cost measures, cost (% Avg.) and cost (% LB.). Only a few correlations were found to be significant with these measures, but because of a weak correlation coefficient, these were practically insignificant. Therefore, $H5_0$ was accepted; that no direct relationship exists between cost criterion and other qualifications criteria.

7.4 CONCLUSIONS

As design professionals and clients are in constant contention over the induction of cost as a selection criterion, it has become much more important to determine how cost incorporation affects the procurement process. One hundred and twenty-two design projects procured using BV and QBS method were analyzed to determine characteristics of the selected bidders, the level of differentiation achieved by different evaluation criteria, and cost criterion relationships with qualifications criteria. All projects analyzed in the sample were public projects from various state, city, local, municipality levels, and university. An analysis file was developed which had procurement evaluation scores of the 122 projects. The analysis file contained scores for each project and each bidder in that project for the six evaluation criteria: cost proposal, schedule proposal, technical proposals, interview, PPI, and RE. Descriptive analysis using frequency distribution, matrix, and measure of central tendency was employed when describing the characteristics of selected bidders. Kruskal-Wallis H tests and Spearman's correlation coefficients were used as inferential tests in order to determine the level of differentiation achieved by the six evaluation criteria and the relationship of cost with qualifications criteria.

In BV projects, more than half the time, the best qualification was selected, compared to 100% in a QBS system. Also, 48 % of the time, the lowest bidder was selected. Lastly, the selected bidder was on average placed first in interview and RE, second in cost, schedule, and technical proposals, and third in PPI.

Moreover, out of the 48% of the time, when lowest bidder was selected, 7% of which, the owner selected the lowest bidder by will, due to the best scoring bidder being too high. Therefore, it was concluded that the cost criterion does influence the owners' selection process and have a disproportionate effect on selection outcome.

When determining the level of differentiation, cost proposal, schedule proposal, and interview attained the greatest differentiation, while technical proposals showed a moderate differentiation. PPI and RE gave the least differentiation. $H1_1$ was therefore accepted, that different evaluation criteria will result in various levels of differentiation between competing A/E firms.

Further analysis of the data sample showed that in almost every project characteristic, at least one evaluation criteria was found to be significant amongst the sub-factors. Hence, $H2_1$ was accepted, that the evaluation criteria will achieve various levels of differentiation based upon separate project characteristics.

Furthermore, the effect the owners' provided project budget and schedule information in the REQ/RFP had on the submitted cost and schedule proposals yielded a significant result for cost and schedule median COV values. This led to acceptance of $H3_1$ and $H4_1$, that the provided project budget and schedule results in lesser variation in submitted cost and schedule proposal by the bidders. When cost criterion was analyzed for relationships with other qualifications criteria, the results showed that cost criterion had no practical direct relationship with any of the qualifications criteria and hence, $H5_0$ was accepted, that no direct proportional relationship exists between the cost criteria and other qualifications criteria.

7.5 CONTRIBUTIONS

7.5.1 CONTRIBUTIONS TO THE BODY OF KNOWLEDGE

Previous research has not assessed the outcomes of best value procurement of A/E consultants, particularly based upon industry-recommended guidelines of a two-envelope method, where the cost proposal is to be held separate from the qualifications, and limitation of cost weights to 30% or less. In the past, researchers have only focused on different procurement models for BV procurement, design professionals' performance and pre-qualification of design professionals. No study has used an actual procurement data, rather had only analyzed different RFQ/RFPs. The methodology employed for analyzing the projects in thesis is novel where the evaluation team scores of bidders for each project were investigated.

Furthermore, no research has yet analyzed the procurement of design professionals under industry wide recommended guide lines of two-envelope method, where the cost proposal is to be held separate from the qualifications and cost weights are to be limited to 30% or less.

Studies that were well known and found to be the most relevant, Chinowsky and Kinsgley (2009) and Christodolou et.al (2004), had the lower sample size of and were limited to smaller geographic. This study analyzes relatively a larger sample size of 122 design professional projects across North America, which analyzed the selected bidders' characteristics, level of differentiation in scores, and cost relationships with qualifications criteria.

Lastly, this study not only analyzes the evaluation team scores of the selected bidders, but instead all the 804 proposing consultants of the 122 projects.

7.5.2 CONTRIBUTIONS TO INDUSTRY PRACTITIONERS

Greater qualifications do not correspond with higher design costs/fees. A common criticism of traditional QBS is that it does not enable the owner to truly check market pricing because they only receive a fee proposal from the single A/E firm that is selected. This leads to fears that the selected A/E firm may not be submitting the most competitive price/cost/fee, whether due to lack of competition or concerns that the highest qualified firm may not be the most cost effective. Yet results of this study show that when owners require a cost/price/fee proposal – in a truly competitive bid environment – there is no correlation between more qualified firms submitting higher cost proposals. In short, QBS is exposed to the perception that hiring the greatest qualifications may cost the owner more and analysis of competitive fee proposals in this study contradicts this perception.

The outcomes of best value procurement for A/E services closely mirrors the results of traditional QBS when cost/price/fee proposals are properly weighted and evaluated. Best value procurement methods resulted in the selection of the most qualified A/E firm in nearly half of cases. In fact, the highest qualified A/E firm also coincided with the lowest cost/price/fee proposal in a quarter of cases. This result indicates that best value procurement outcomes are both cost-effective and also closely mirror QBS. Yet in order to achieve these outcomes, Owners must be careful to structure their best value procurement methods within recommended industry guidelines for A/E selections, namely via the use of a two-envelope evaluation system where cost/price/fee proposals do not account for more than thirty percent of the total evaluation score.

Forward-looking, project-specific evaluation criteria were found to more effectively differentiate between competing A/E qualifications than criteria such as past performance or related experience. Interviews with the A/E project team resulted in the largest variance between competing bidders (COV=17.1%) of the qualifications-based evaluation criteria included within this study. Technical proposals were second and had a coefficient of variation of nearly 15%. Other qualifications-based criteria of past performance and related, conversely, were each found to achieve little-to-no differentiation (4.8% and 0.0%, respectively). These results show that there is a large disparity in the effectiveness in the various methods that owners use to evaluate the qualifications of A/E firms. Owners who evaluate qualifications in the "rearview" mirror – that is, by focusing on a design team's previous results – tend to find marginal differences compared with owners who emphasize more forward-looking qualifications-criteria. Based on these results, owners are recommended to focus their evaluation of A/E qualifications on the quality of the project team individuals that would be assigned to their project along with the technical proposal of the project-specific means, methods, and innovations the team would employ.

7.6 LIMITATION AND RECOMMENDATIONS FOR FUTURE RESEARCH

Despite the above stated conclusions and contributions, there are a few limitations to this research. The study only analyzed the procurement process and results for design projects, hence no contractor selection or design-build team selection was evaluated. Also, this study only focuses on two-envelope BV and QBS procurement methods and does not cover any other procurement methods. Even though the six main evaluation criteria were investigated in this study, different types of owners use different sets of evaluation criteria when procuring design professionals. Lastly, the sample consisted of only public projects, therefore the results of this research do not depict the selection process and trends for privately procured projects.

Future research with different procurement methods, vendor types, evaluation criteria, and owner type can be added to the research area. Furthermore, this study also opens different avenues of research as future researchers can focus on detailed analysis of how the owner-provided budget, cost, and schedule affect the submitted cost and schedule proposals. Different project characteristics such as level of design, delivery method, etc. can also be investigated for differentiation achieved in scores. Key personnel interview scores can be investigated in detail to determine the level of differentiation achieved by different personnel.

REFERENCES

ACEC (American Council of Engineering Companies) (2006). *ACEC Policy Statement Selection of Design Professionals on the Basis of Qualification*. ACEC, Washington, DC.

AIA (The American Institute of Architects) (2011). *Issue Brief Qualification Based Selection*. AIA, Washington, DC.

APWA (American Public Work Association) (2008). *Qualifications Based Selection of Professional Services Consultants*. APWA, Washington, DC.

AWWA (American Water Works Association). (2016). "Qualifications-Based Selection of Professional Services." *Policy Statement*, < http://www.awwa.org/about-us/policy-statements/policy-statement/articleid/4039/qualifications-based-selection-of-professional-services.aspx> (July. 29, 2016).

Cheung, F. K. T., Kuen, J. L. F., and Skitmore, M. (2002). "Multi-criteria evaluation model for the selection of architectural consultants." *Construction Management and Economics*, 20(7), 569-580.

Chinowsky, P. S., and Kingsley, G. A. (2009). "An analysis of issues pertaining to qualifications-based selection." American Council of Engineering Companies and American Public Works Association, Washington, DC.

Christodoulou, S., Griffis, F., Barrett, L., and Okungbowa, M. (2004). "Qualifications-Based Selection of Professional A/E Services." *J. Manage. Eng.*, 10.1061/(ASCE)0742-597X(2004)20:2(34), 34-41.

Chow, L, K., and Ng, S, T. (2007a). "Expectation of performance levels pertinent to consultant performance evaluation." *International Journal of Project Management,* 25(1), 90-103.

Chow, L, K., and Ng, S, T. (2007b). “A fuzzy gap analysis model for evaluating the performance of engineering consultants.” *Automation in construction*, 16(4), 425-435.

Feldmann, M., Chrusciel, D., Pohlmann, A., Shelley, M., II, McCool, K., Morton, A., and Ahoy, C. (2008). "Architectural and Engineering Fees from the Public Institutional Perspective." *J. Manage. Eng.*, 10.1061/(ASCE)0742-597X(2008)24:1(2), 2-11.

Fleenor, C., and Hall, S. (2002). “The state of QBS: battles at the local and national level are challenging how firms do business.” *Engineering Inc.,* 13(5), 16–19.

Hampton, D. (1994). "Procurement Issues." *J. Manage. Eng.*, 10.1061/(ASCE)9742-597X(1994)10:6(45), 45-49.

Horns, K., and Jenkins, R. (2011). "Is the Profession of Civil Engineering Becoming a Commodity? You Should Know the Answer." *Leadership Manage. Eng.*, 10.1061/(ASCE)LM.1943-5630.0000089, 40-44.

Ling, Y, Y. (2002). "Model for Predicting Performance of Architects and Engineers." *J. Constr. Eng. Manage.*, 10.1061/(ASCE)0733-9364(2002)128:5(446), 446-455.

Ling, Y, Y. (2003). “A conceptual model for selection of architects by project managers in Singapore.” *International Journal of Project Management*, 21(2), 135-144.

Nguyen, T. H., and Shehab, T. (2008). “Selecting an architecture-engineering team by using fuzzy set theory.” *Engineering, Construction and Architectural Management*, 15(3), 282-298.

NSPE (National Society of Professional Engineers). (2016). “Qualifications-Based Selection of Engineering Services.” *Issues and Advocacy*, < https://ww.nspe.org/resources/issues and-advocacy/action-issues/qualifications-based-selection-engineering-services> (July. 29, 2016).

Oyedele, L, O., and Tham, K, W. (2007). “Clients’ assessment of architects’ performance in building delivery process: Evidence from Nigeria.” *Building and Environment,* 42(5), 2090-2099.

Parks, G. (2006). "At a Crossroads— Is Value-based Compensation the Answer?" *Leadership Manage. Eng.*, 10.1061/(ASCE)1532-6748(2006)6:4(144), 144-149.

Sullivan, K, T., Savicky, J., and Carey, B. (2010). “Best-Value Process Implementation at the City of Peoria: Five Years of Research Testing.” *Journal for the Advancement of Performance Information and Value*, 2(1), 23-32.

El Wardani, M., Messner, J., and Horman, M. (2006). "Comparing Procurement Methods for Design-Build Projects." *J. Constr. Eng. Manage.*, 10.1061/(ASCE)0733-9364(2006)132:3(230), 230-238.

Thomas Ng, S., and Chow, L. (2004a). "Framework for Evaluating the Performance of Engineering Consultants." *J. Prof. Issues Eng. Educ. Pract.,* 10.1061/(ASCE)1052-3928(2004)130:4(280), 280-288.

Thomas Ng, S., and Chow, L. (2004b). "Evaluating engineering consultants’ general capabilities during the pre-selection process - a Hong Kong study." *Engineering, Construction and Architectural Management*, 11(3), 150-158.

APPENDIX A – BOX PLOTS AND SCATTER PLOTS

Box plot (assumption for Kruskal-Wallis H test):

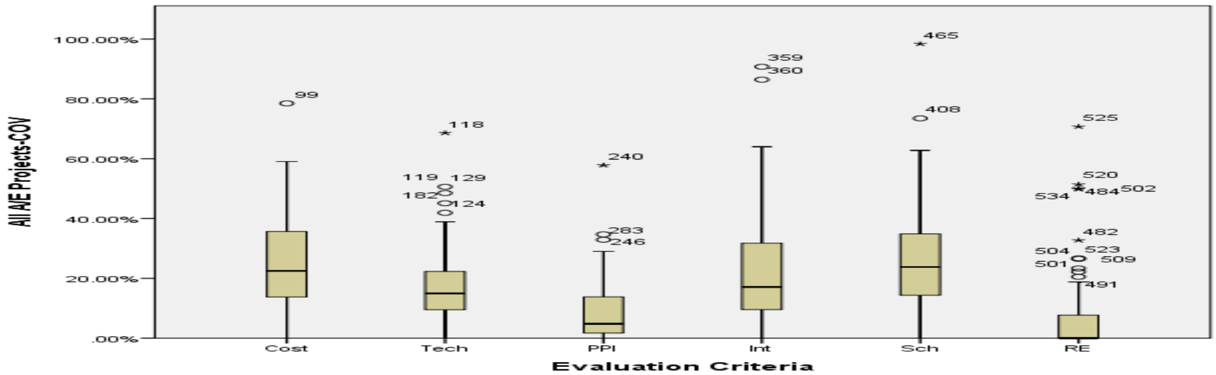

Figure A.1 All Evaluation Criteria for A/E

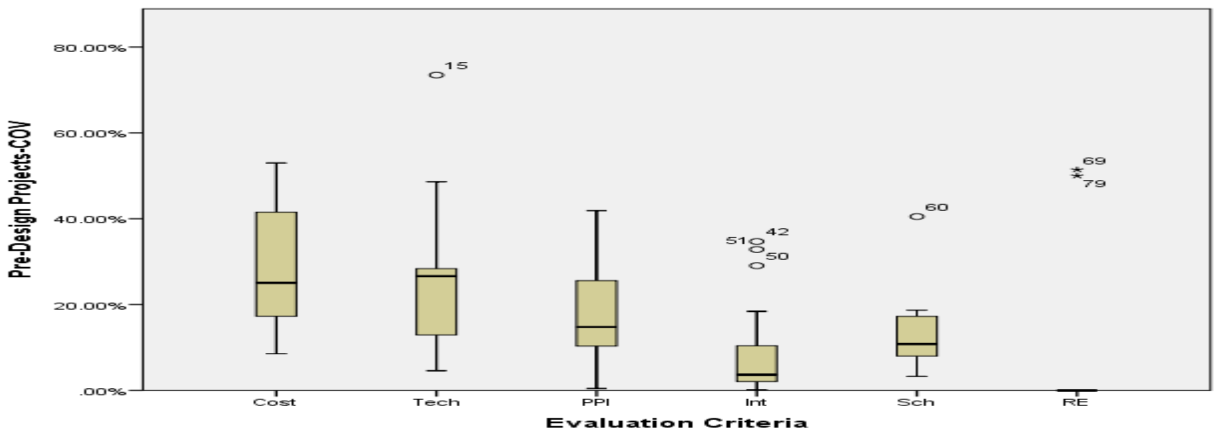

Figure A.2 All Evaluation Criteria for Pre-Design Projects

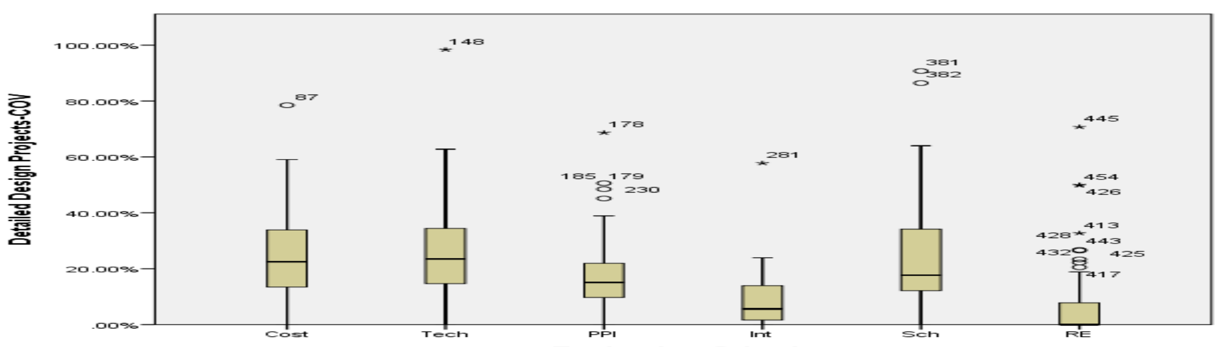

Figure A.3 All Evaluation Criteria for Detailed Design Project

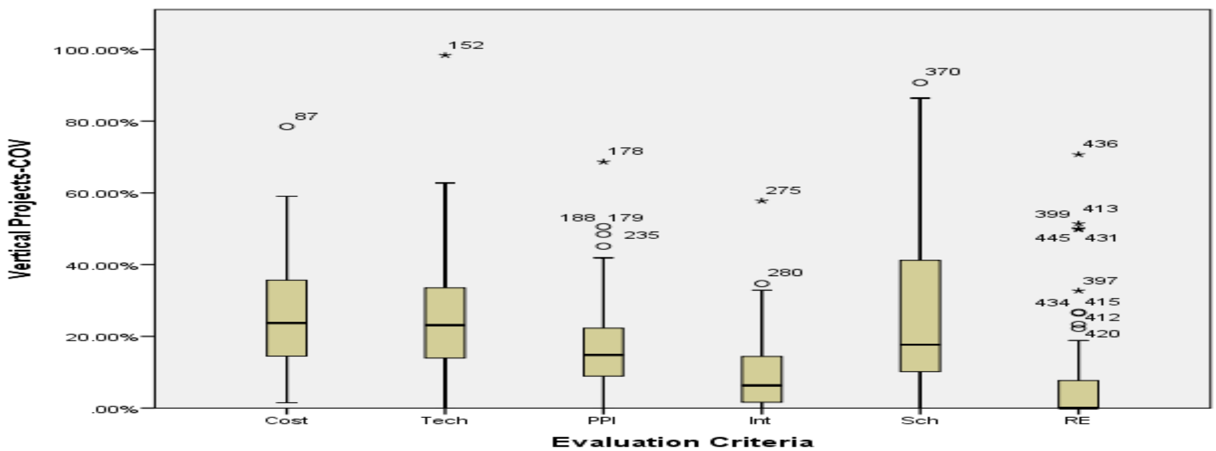

Figure A.4 All Evaluation Criteria for Vertical Project

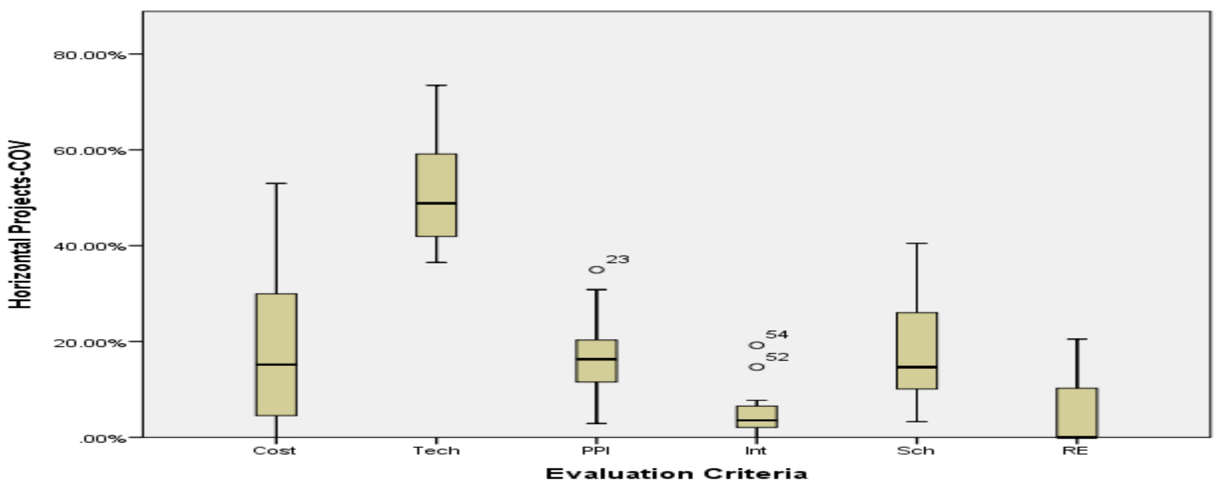

Figure A.5 All Evaluation Criteria for Horizontal Projects

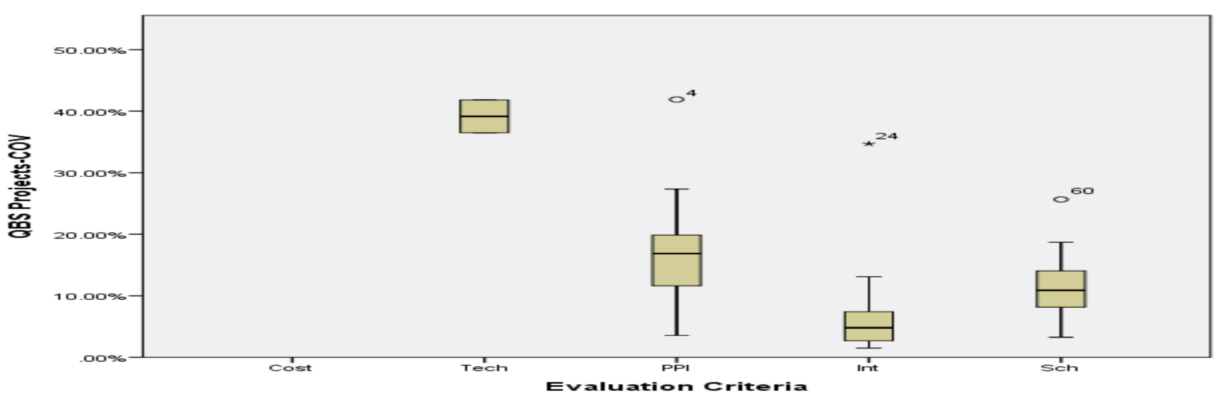

Figure A.6 All Evaluation Criteria for QBS Projects

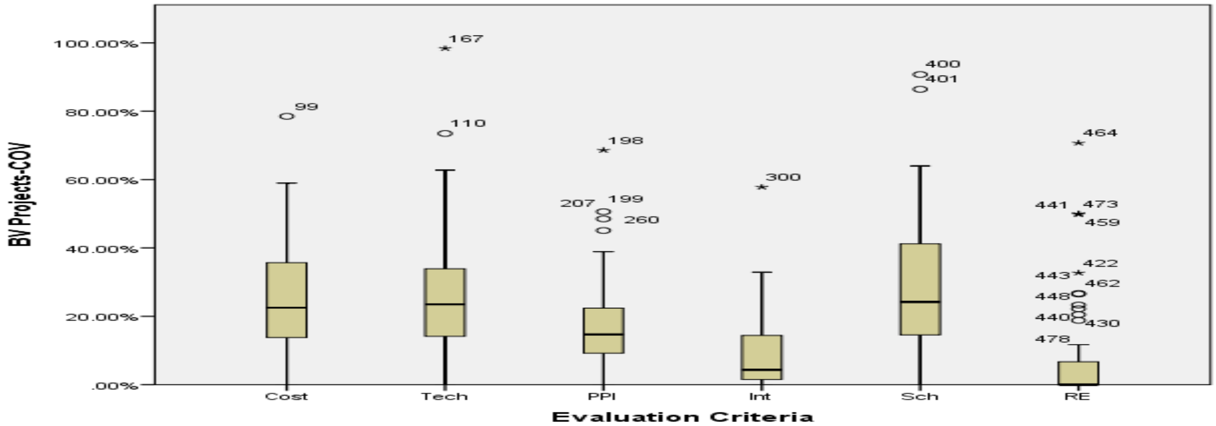

Figure A.7 All Evaluation Criteria for BV Projects

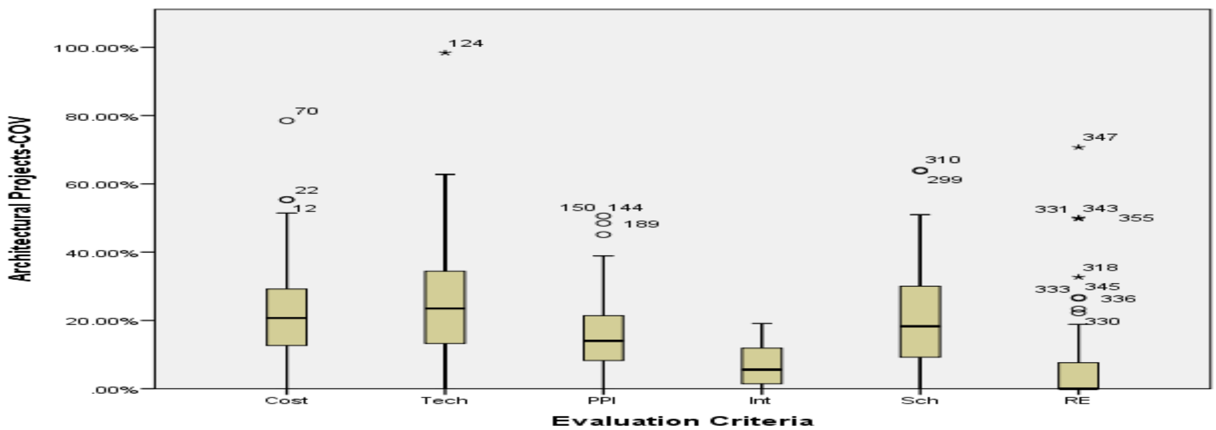

Figure A.8 All Evaluation Criteria for Architectural Projects

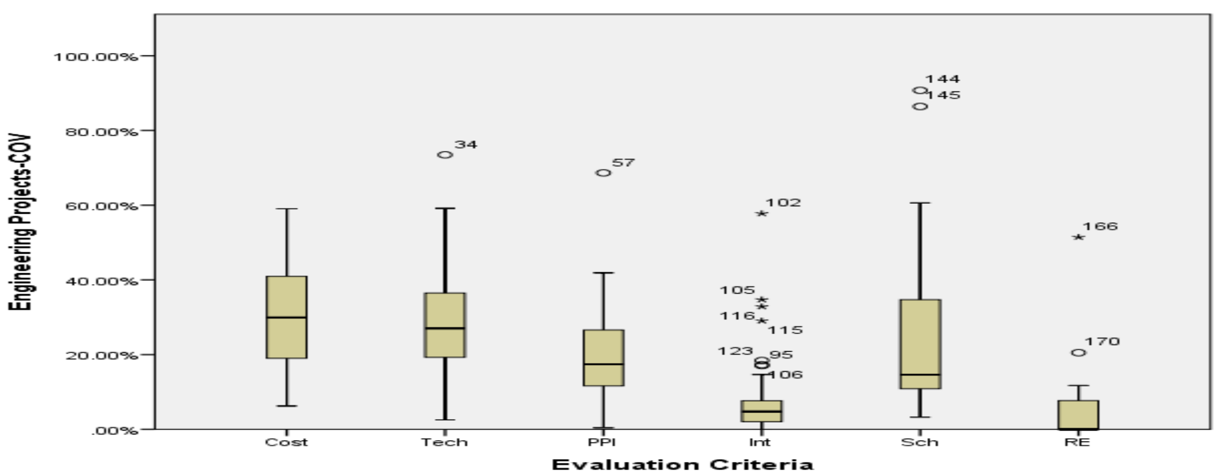

Figure A.9 All Evaluation Criteria for Engineering Projects

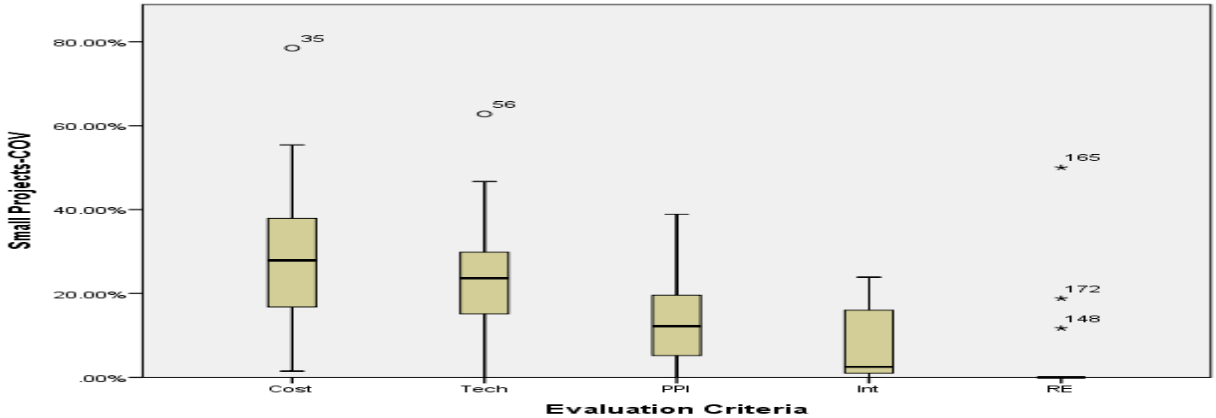

Figure A.10 All Evaluation Criteria for Small Projects

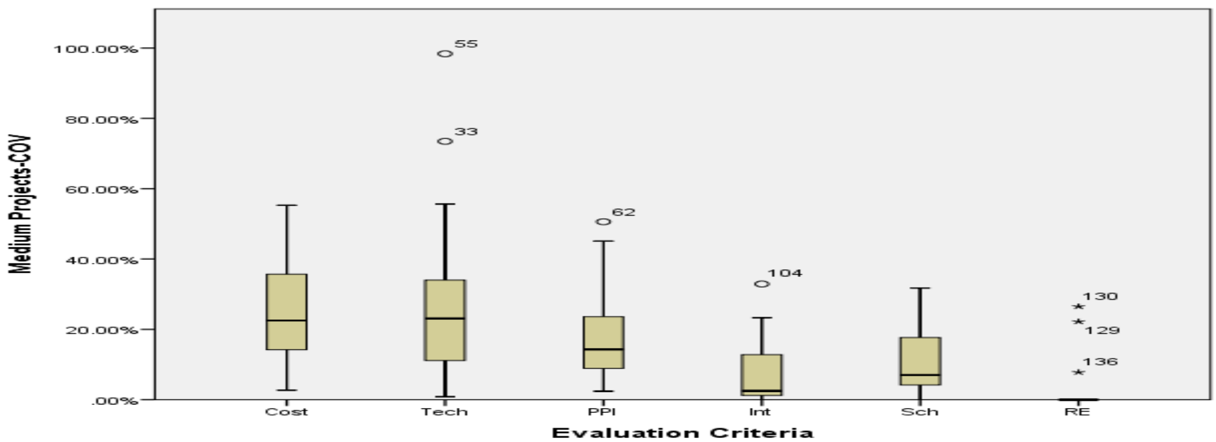

Figure A.11 All Evaluation Criteria for Medium Projects

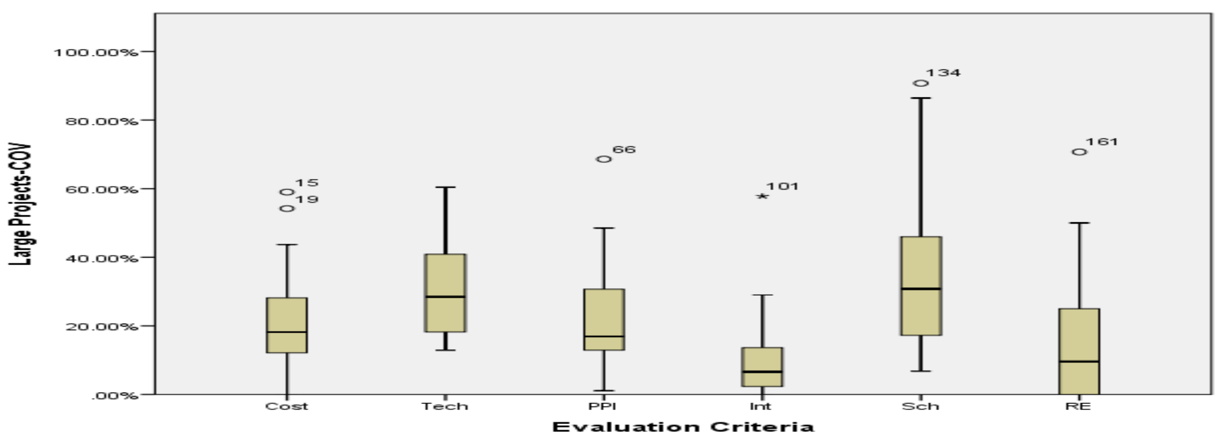

Figure A.12 All Evaluation Criteria for Large Projects

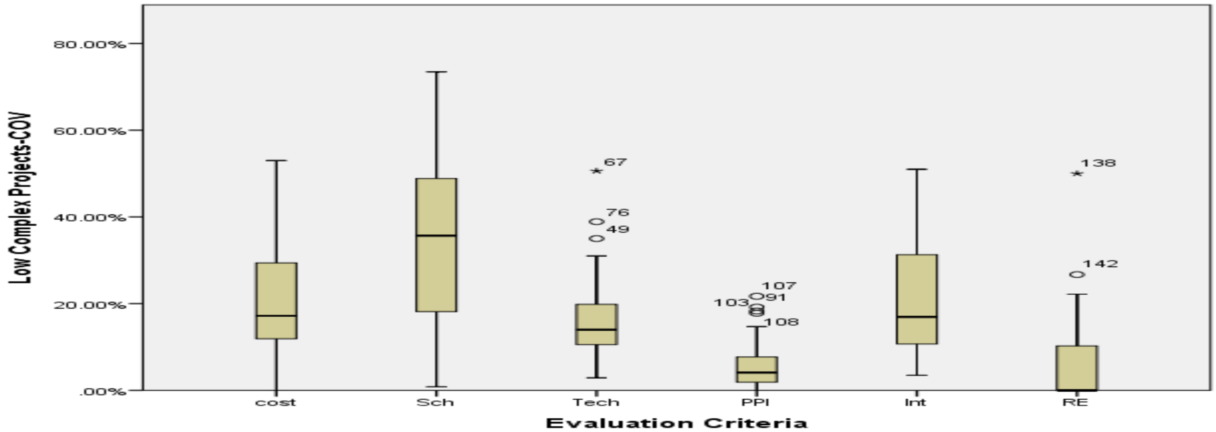

Figure A.13 All Evaluation Criteria for Low Complex Projects

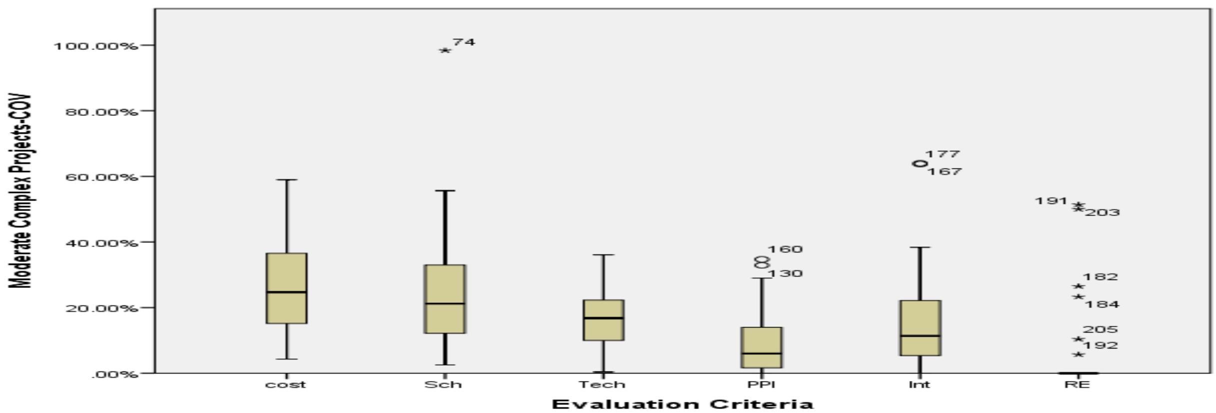

Figure A.14 All Evaluation Criteria for Moderate Complex Projects

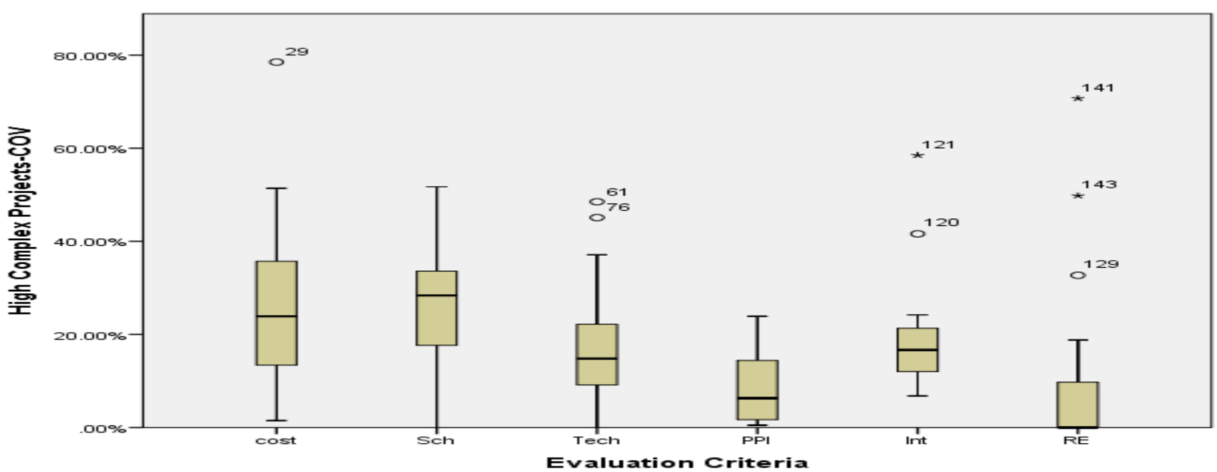

Figure A.15 All Evaluation Criteria for High Complex Projects

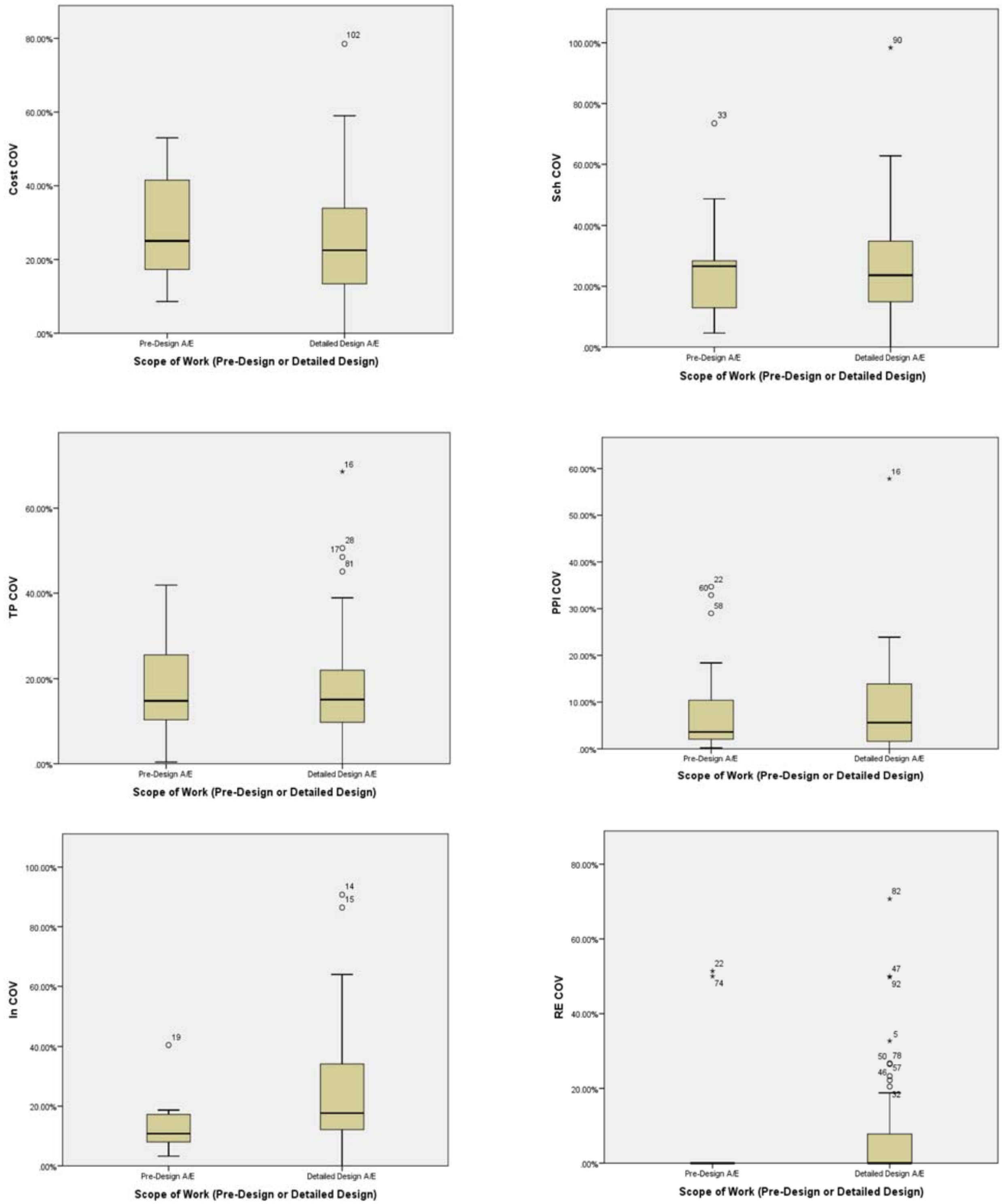

Figure A.16 All Evaluation Criteria for Scope of Work (Pre-Design and Detailed Design

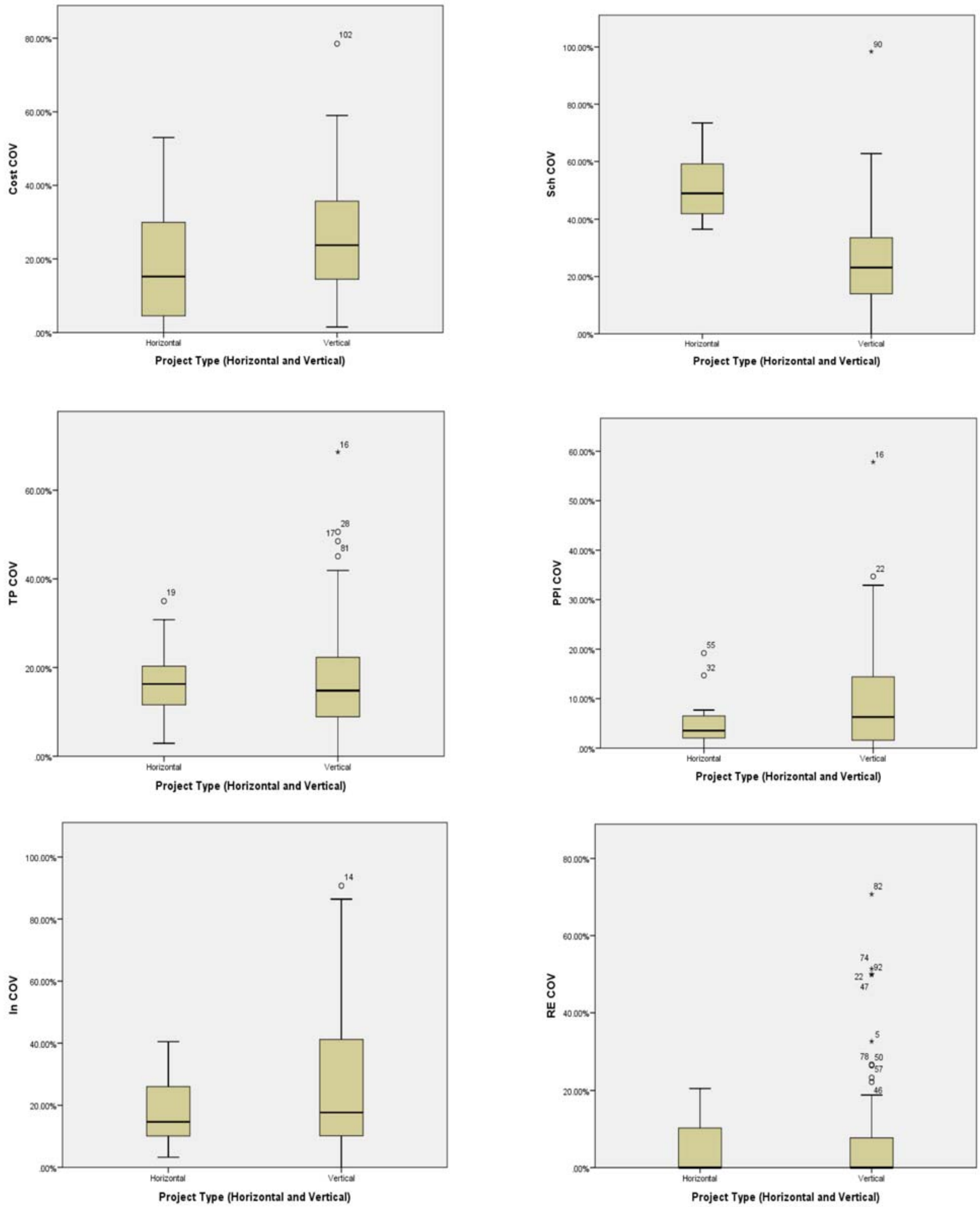

Figure A.17 All Evaluation Criteria for Project Type (Horizontal and Vertical)

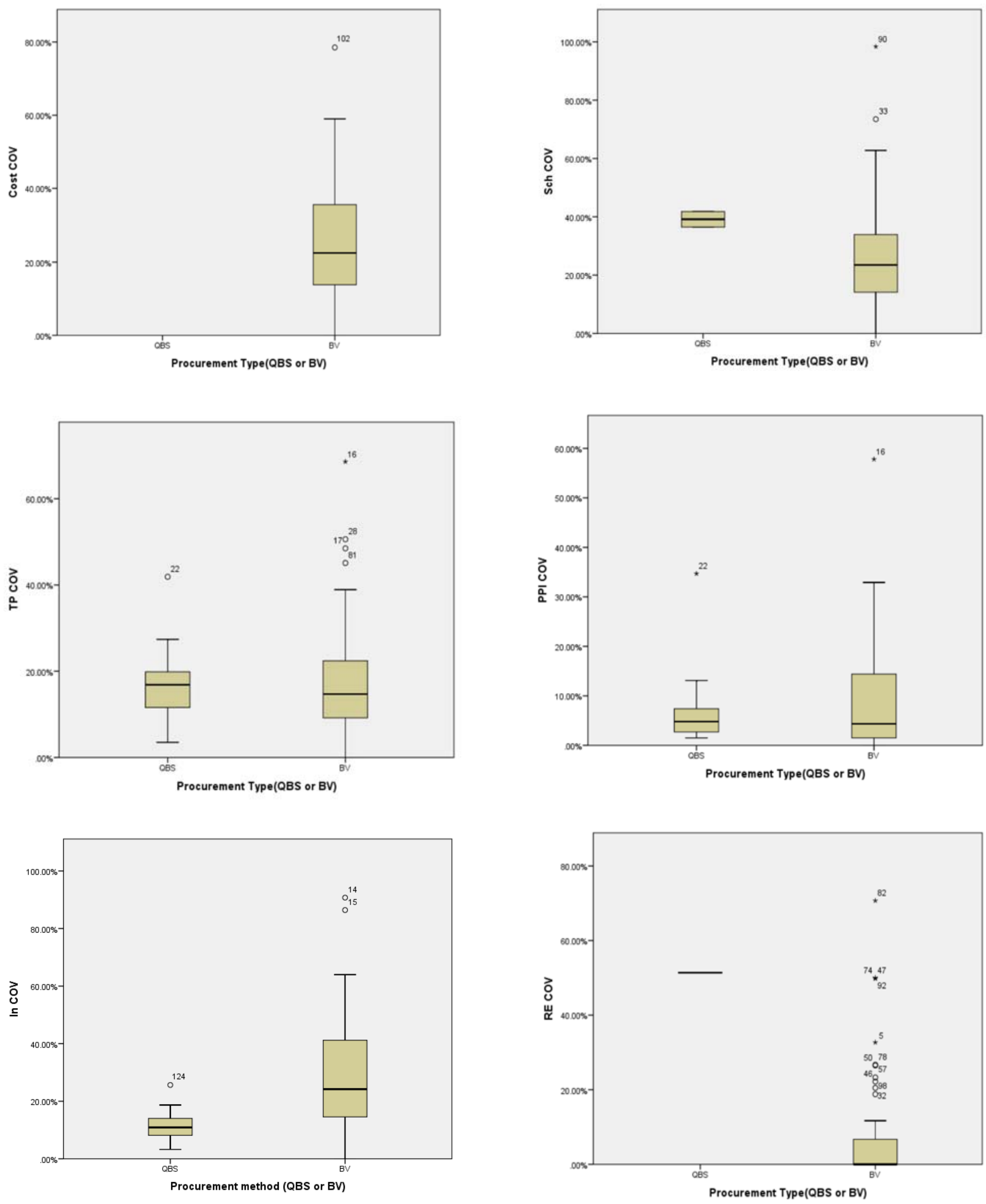

Figure A.18 All Evaluation Criteria for Procurement Type (QBS and BV)

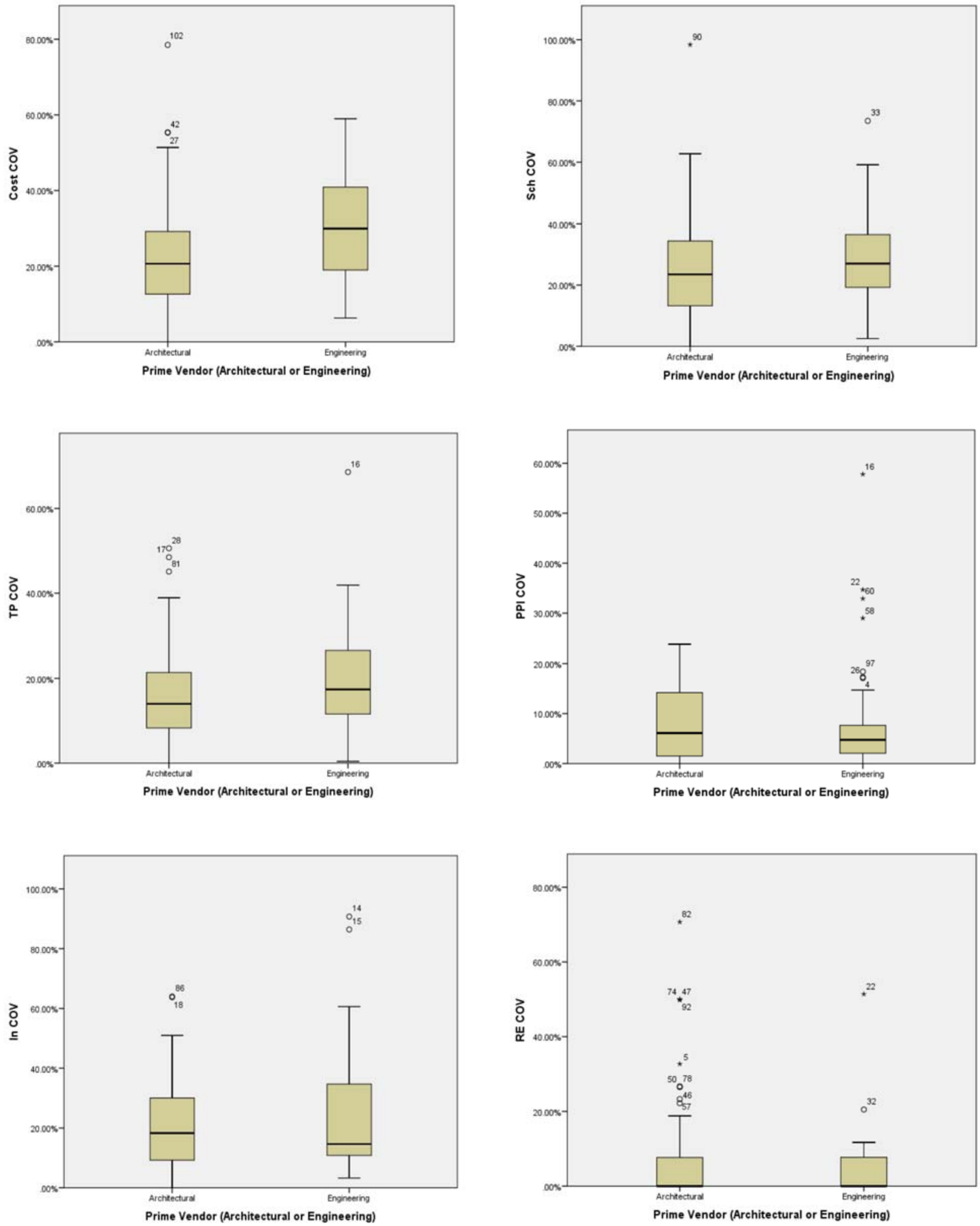

Figure A.19 All Evaluation Criteria for Prime Vendor (Architectural and Engineering)

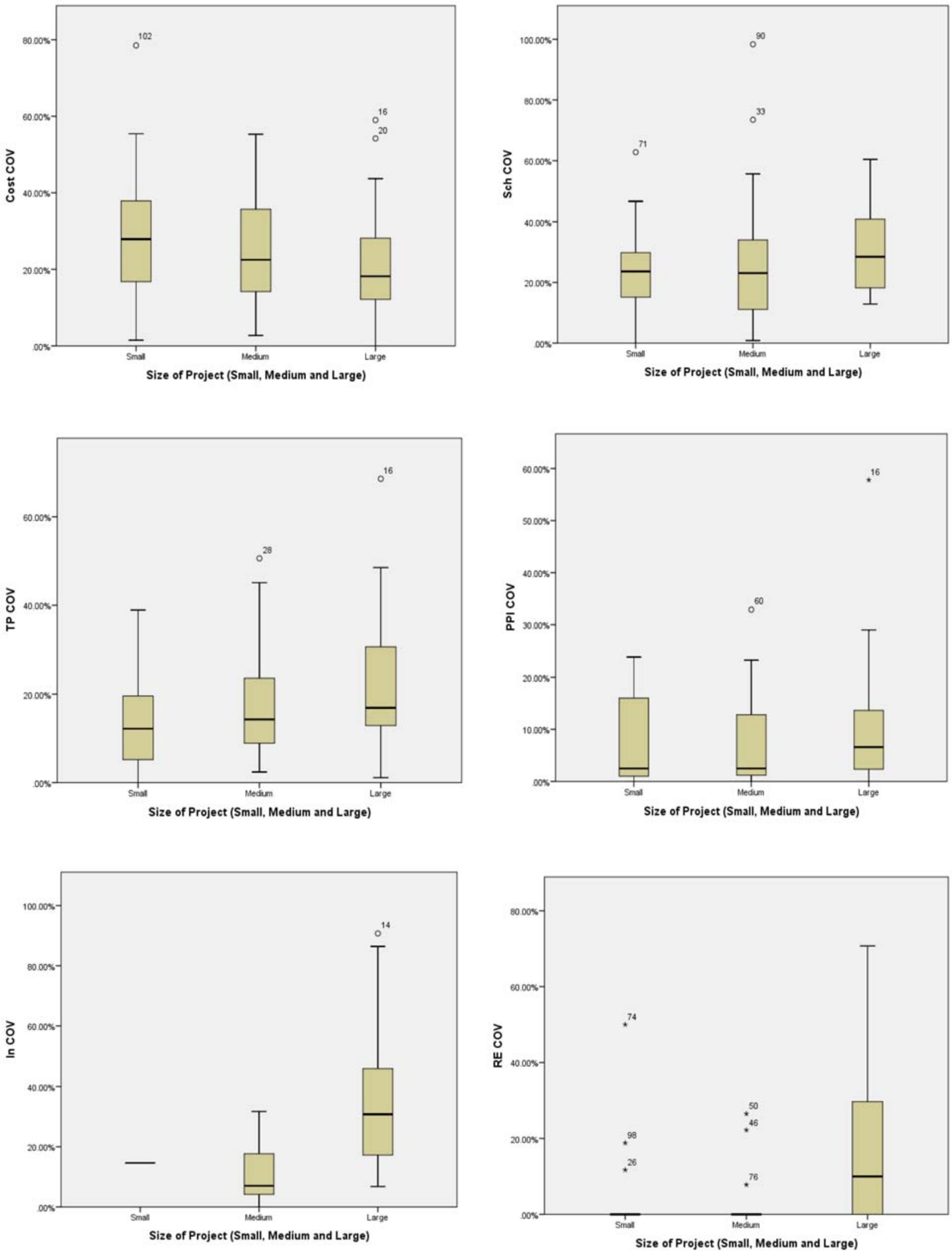

Figure A.20 All Evaluation Criteria for Size of Project (Small, Medium, and Large)

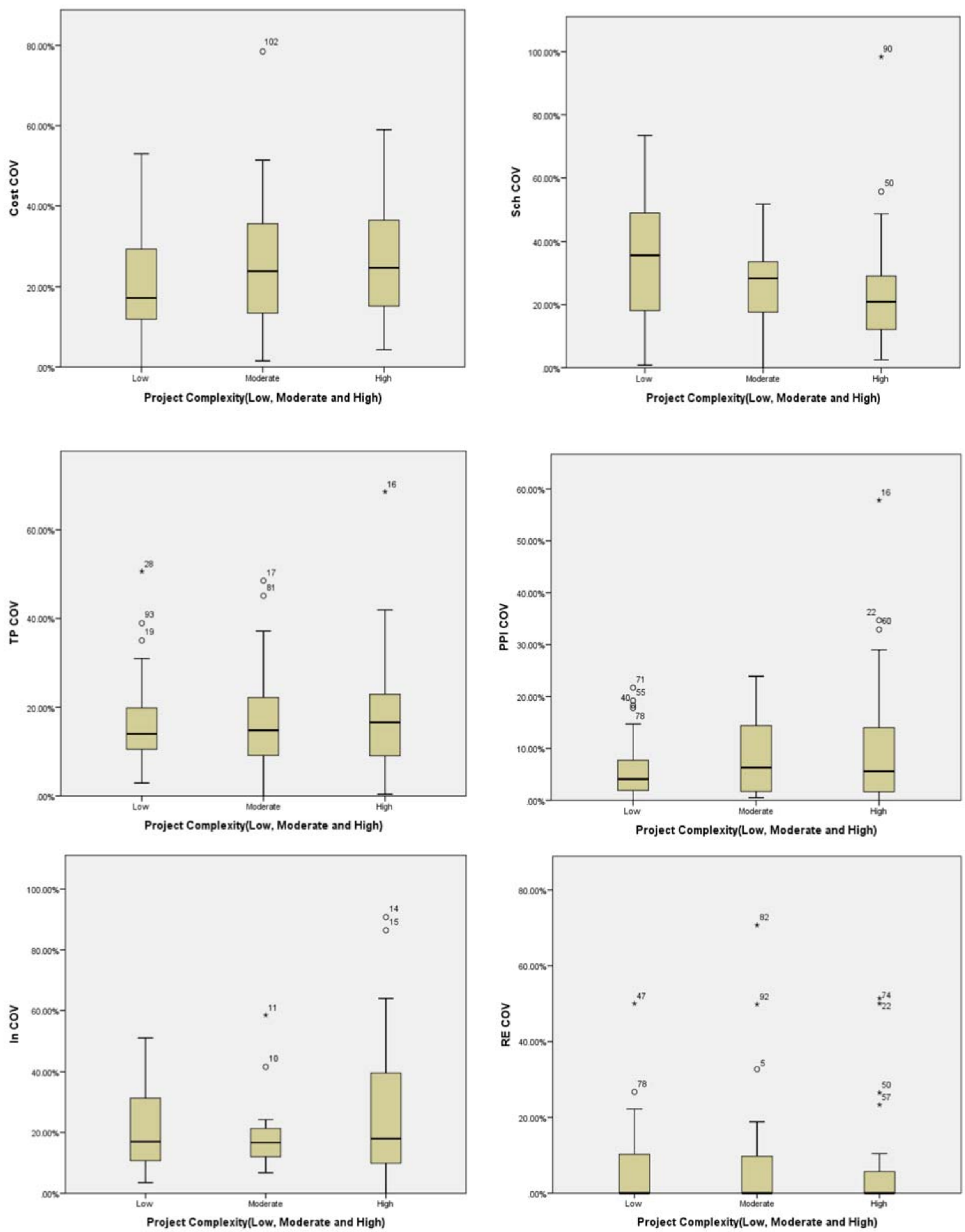

Figure A.21 All Evaluation Criteria for Project Complexity (Low, Moderate and High)

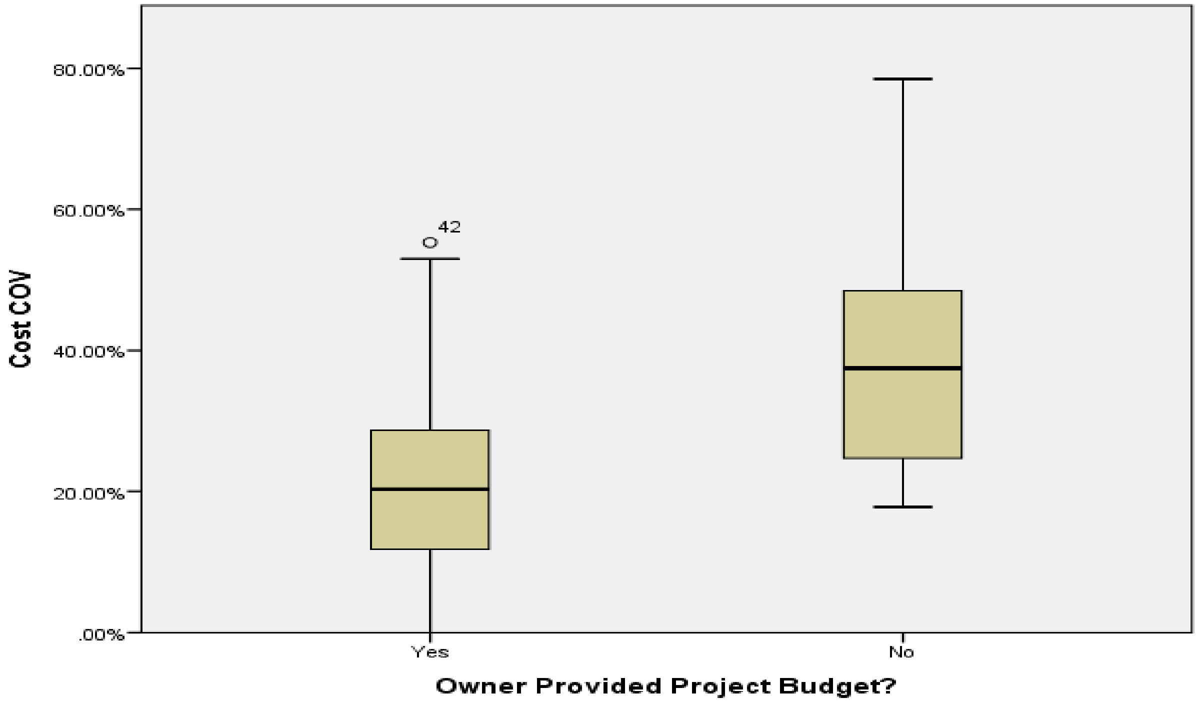

Figure A.22 Cost COV for the Owner Provided Project Budget

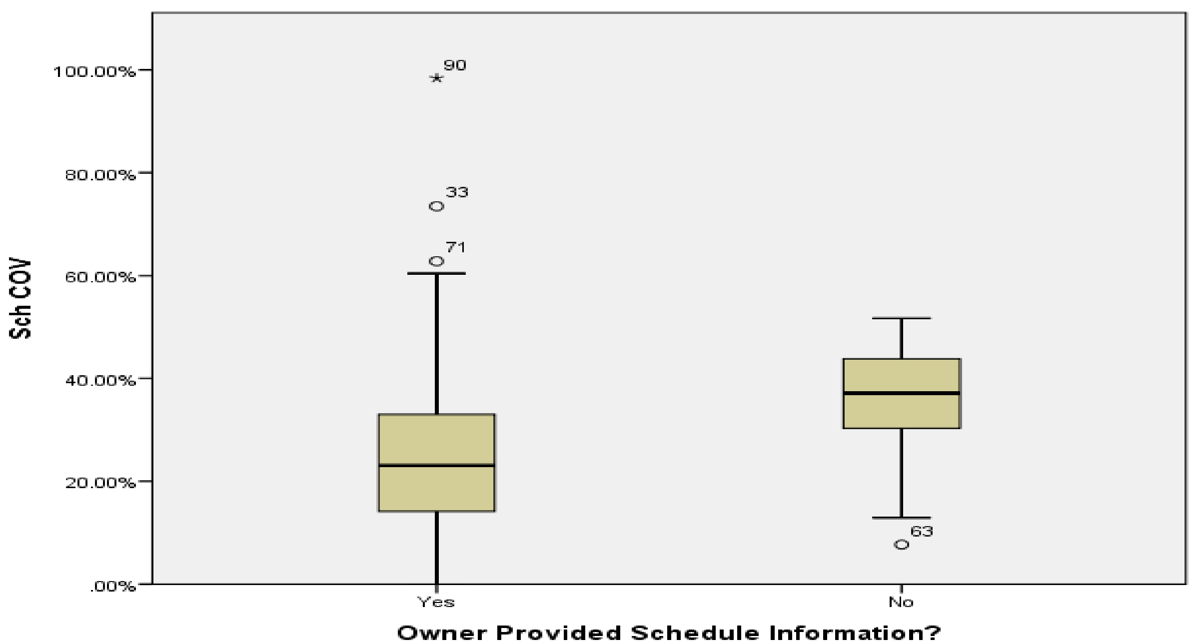

Figure A.23 Schedule COV for the Owner Provided Schedule Information

Scatter plot (monotonic assumption for Spearman’s correlation coefficient):

Cost (% Avg.)

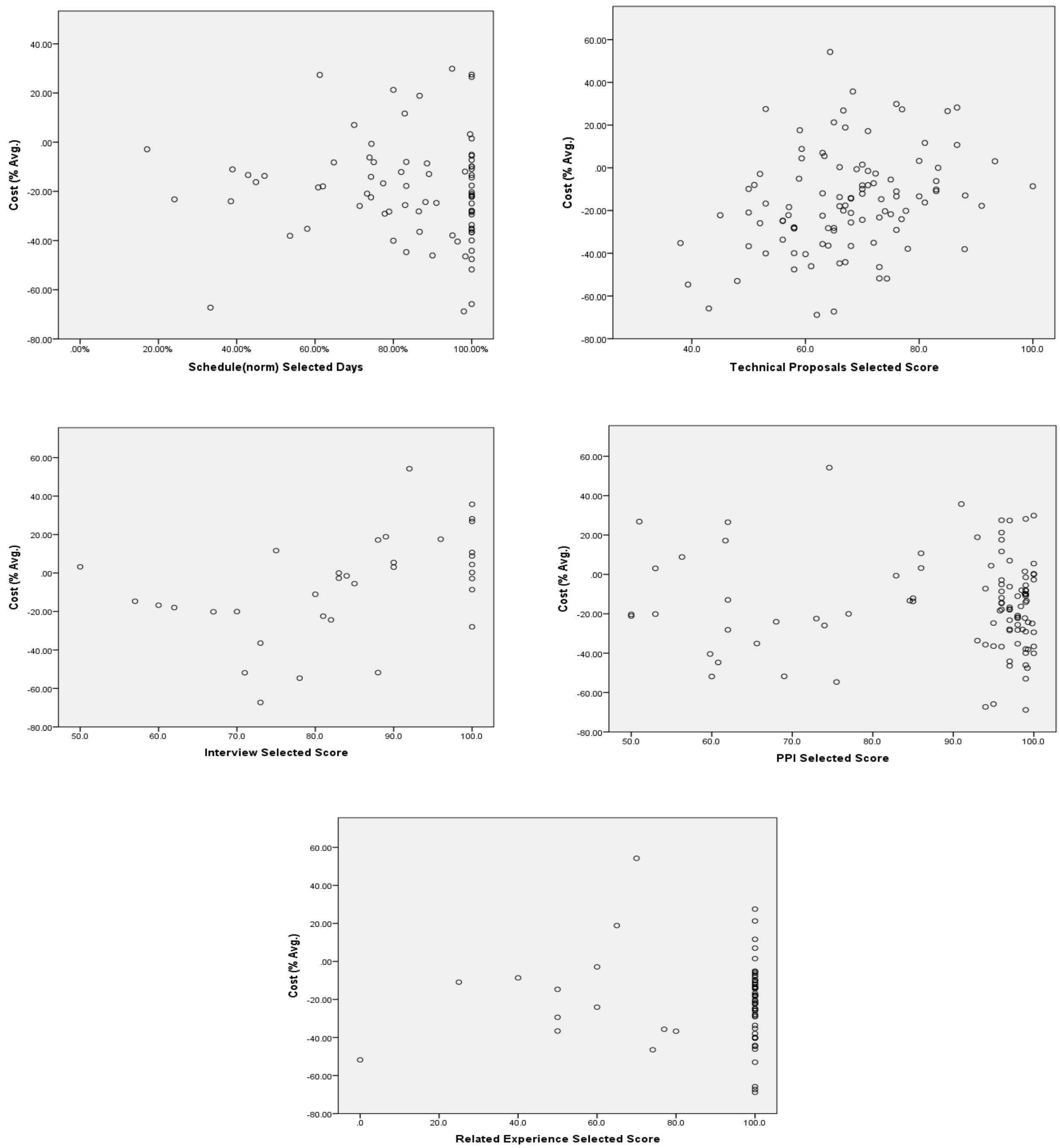

Figure A.24 Scatter Plot for Cost (% Avg.) Vs Evaluation Criteria

Cost (% LB)

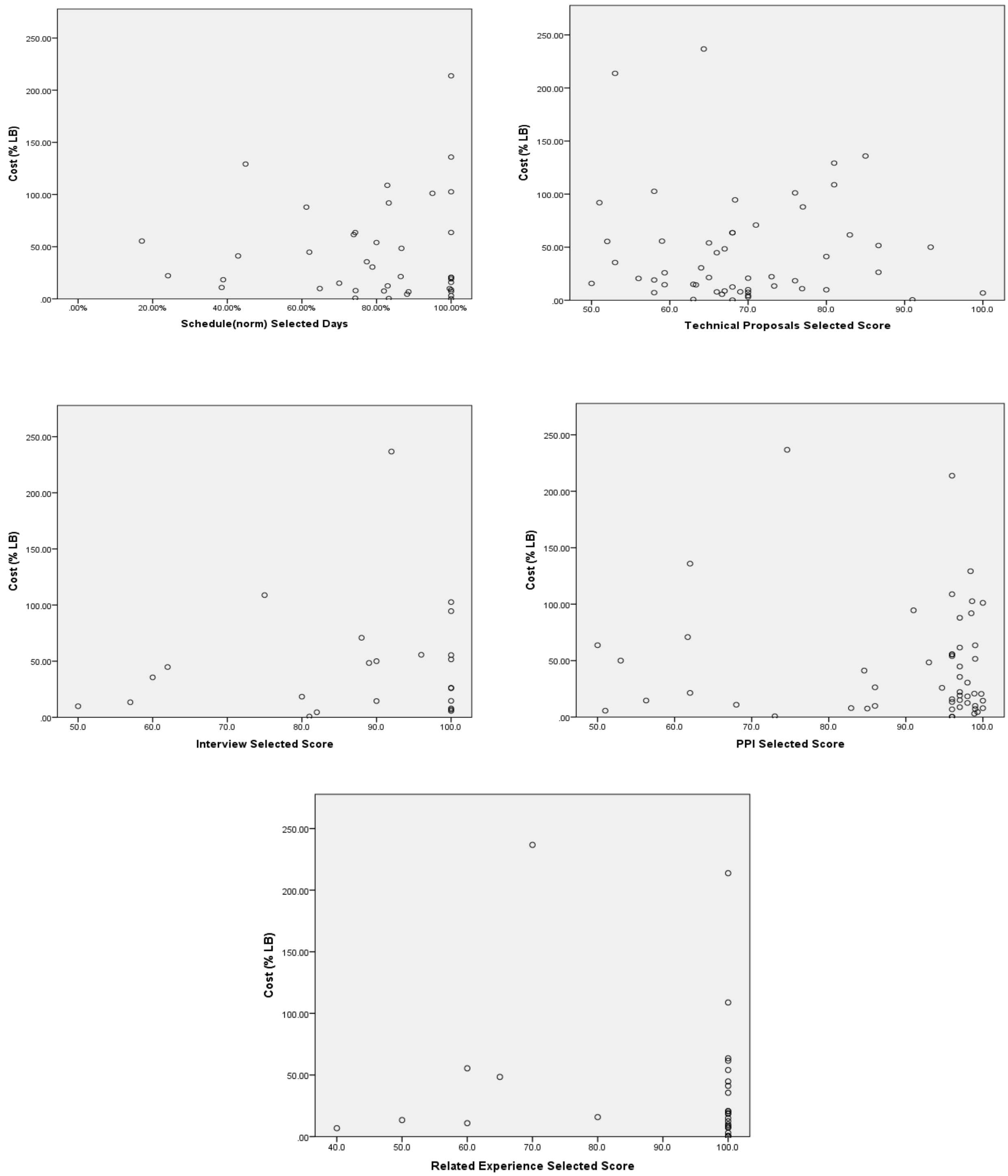

Figure A.25 Scatter Plot for Cost (% LB) Vs Evaluation Criteria

Technical proposals:

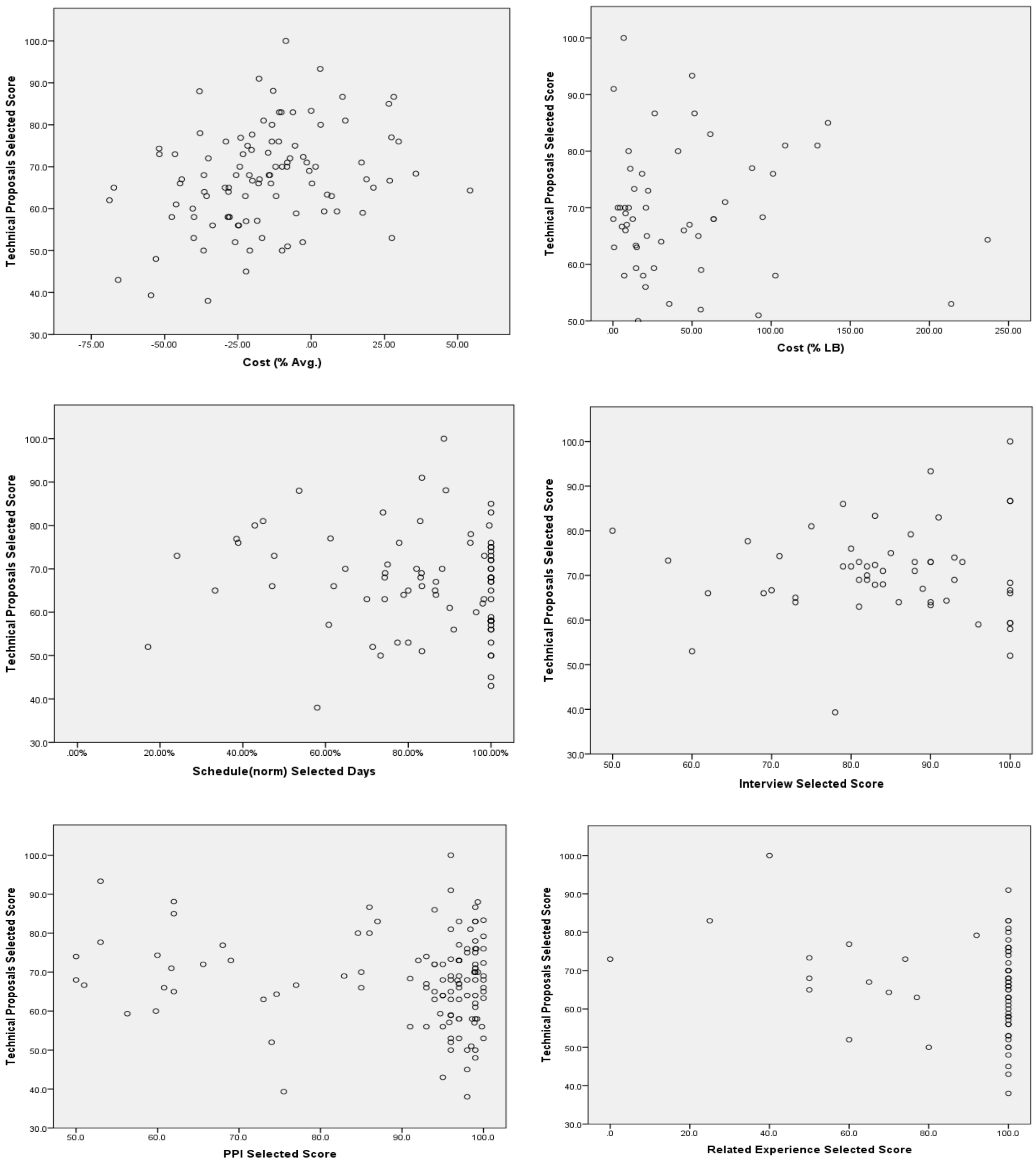

Figure A.26 Scatter Plot for Technical Proposals Vs Evaluation Criteria

PPI:

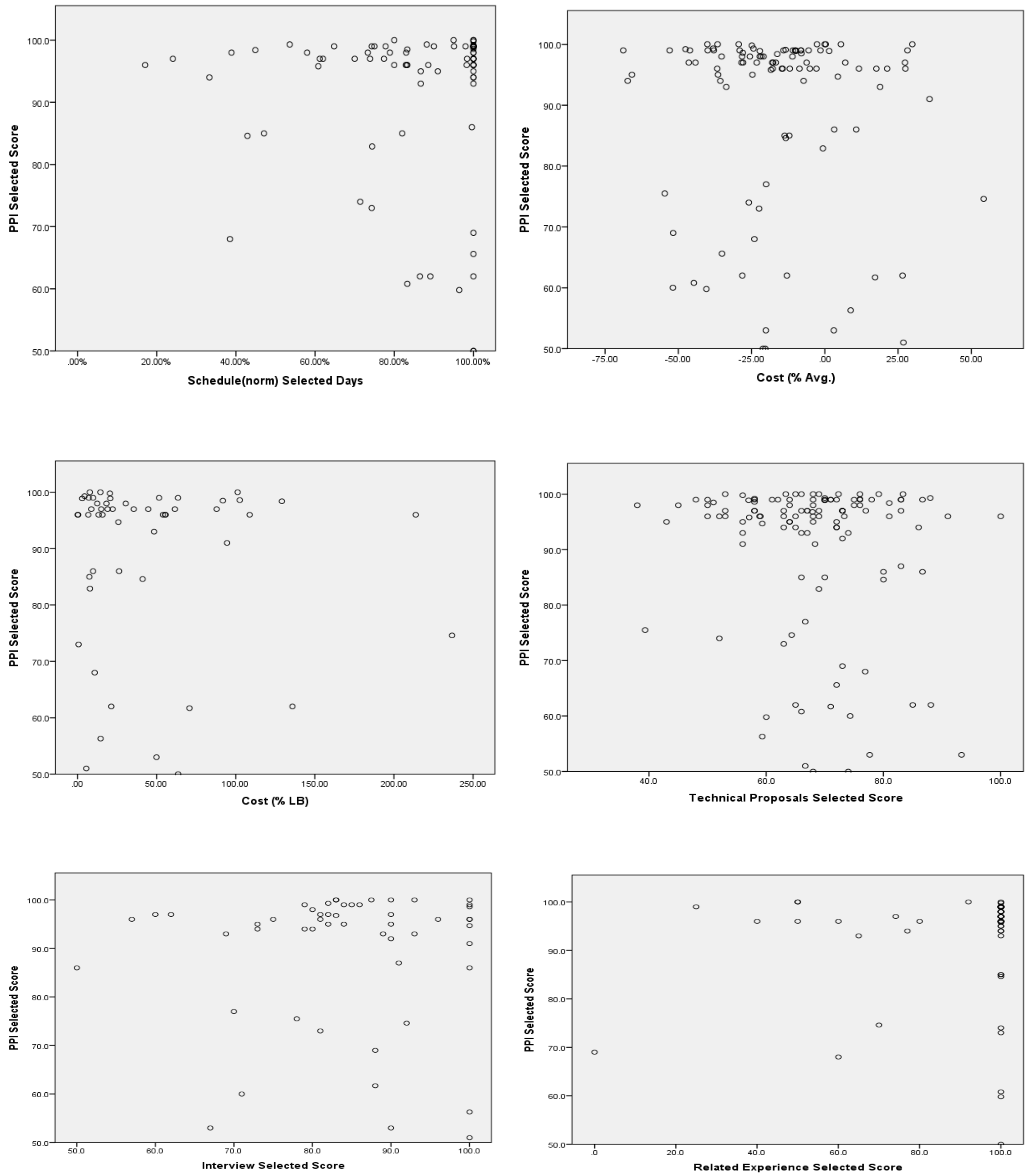

Figure A.27 Scatter Plot for Related Experience Vs Evaluation Criteria

Interview:

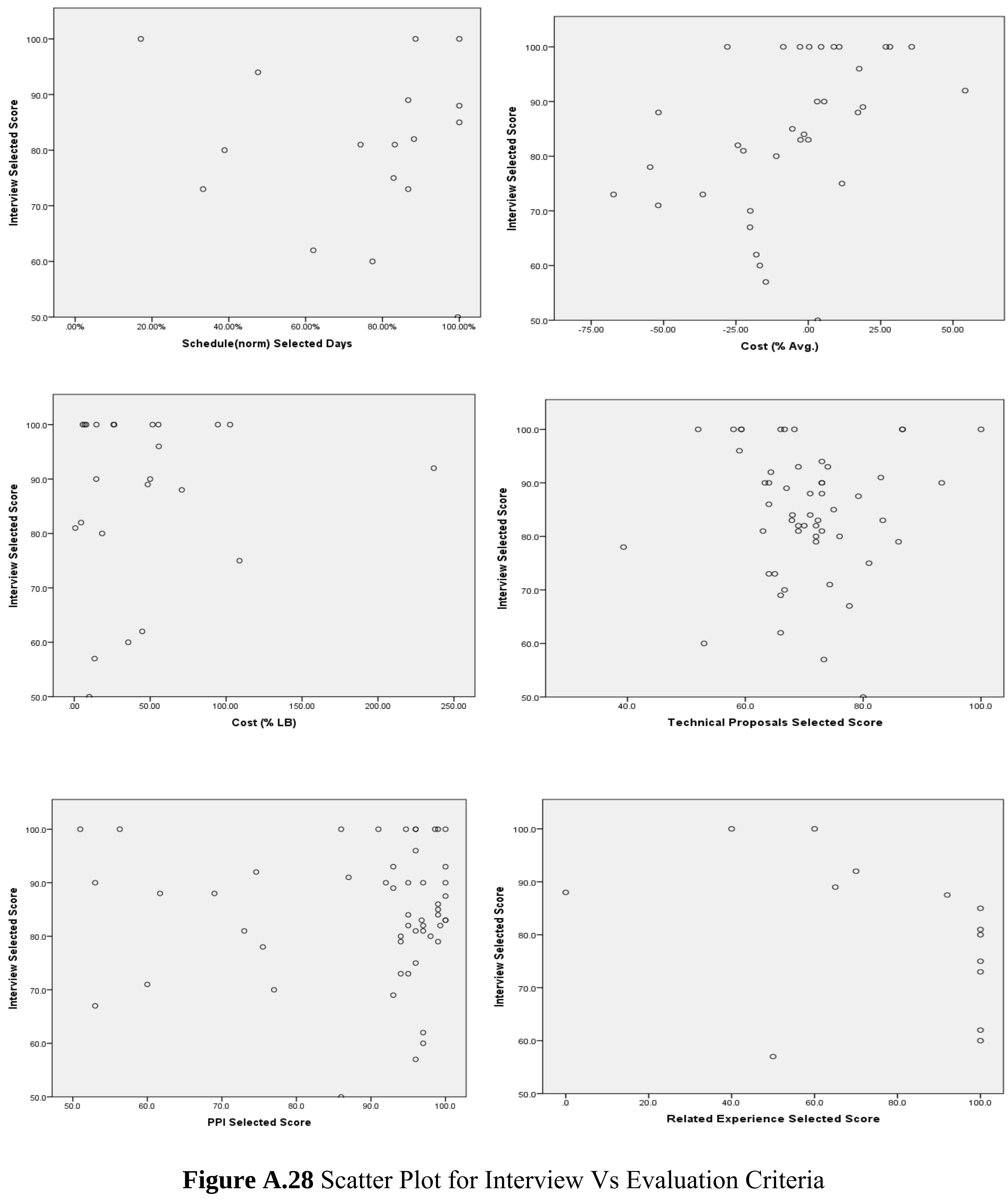

Figure A.28 Scatter Plot for Interview Vs Evaluation Criteria

Related Experience:

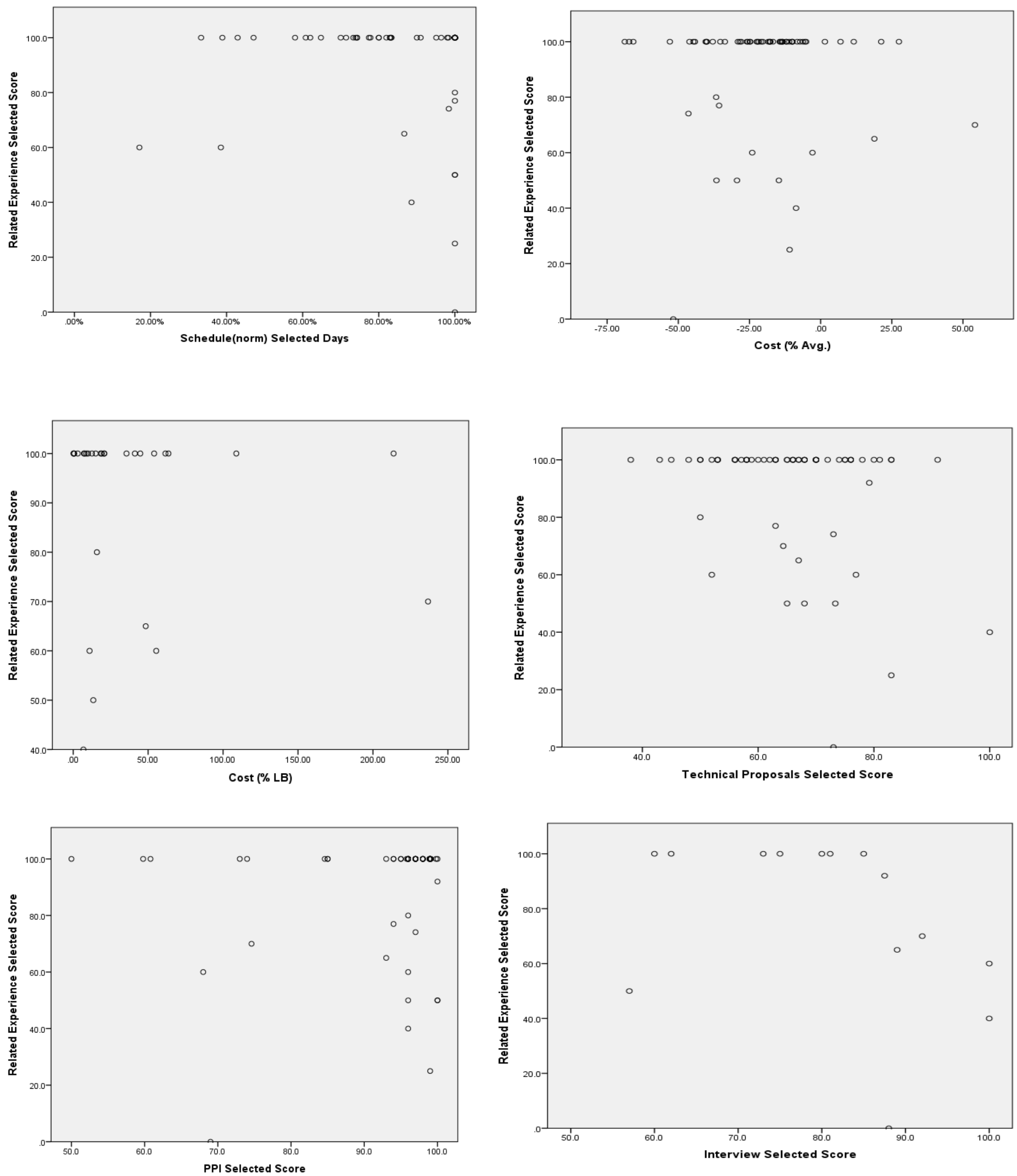

Figure A.29 Scatter Plot for Related Experience Vs Evaluation Criteria

Schedule (norm.):

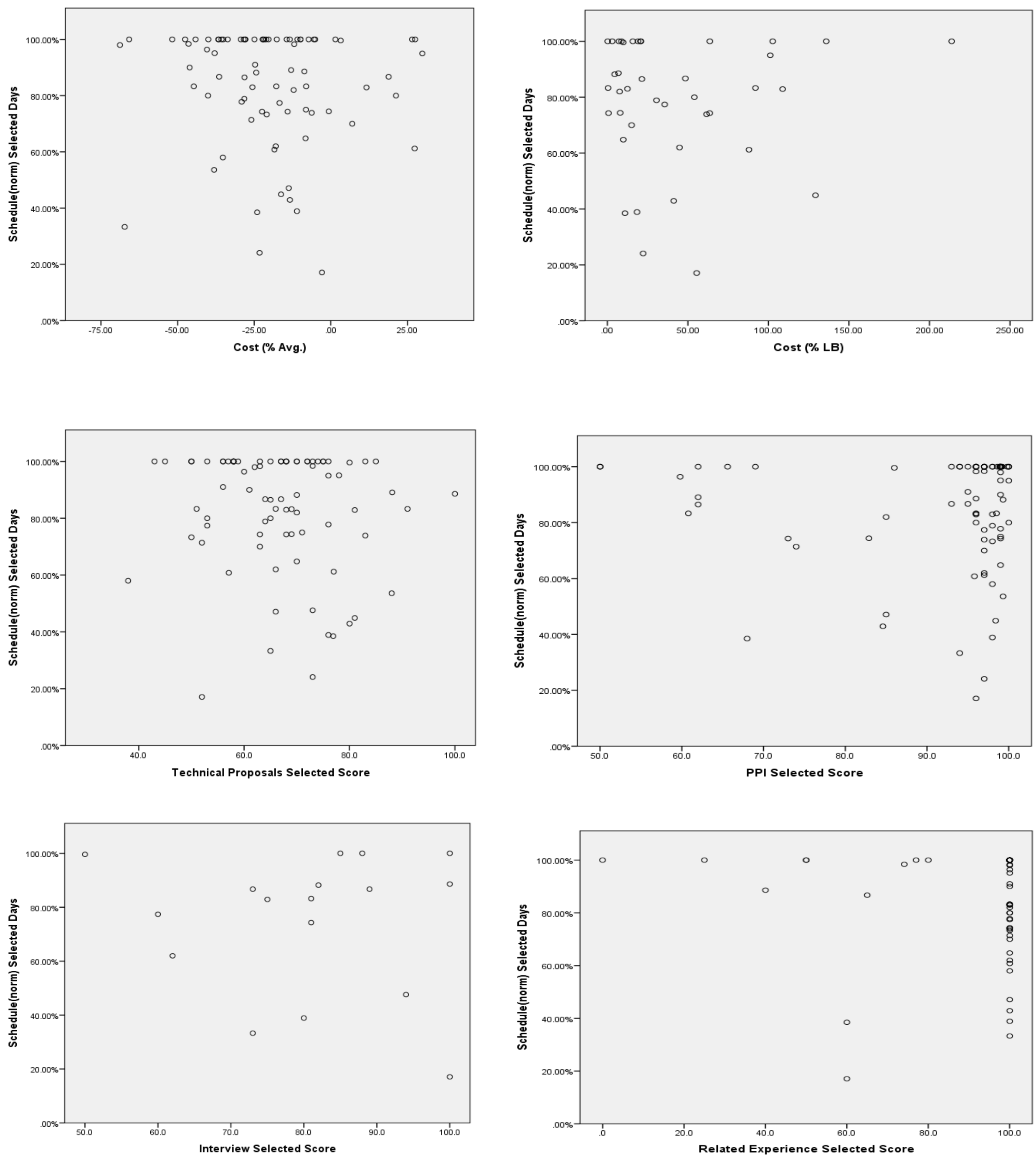

Figure A.30 Scatter Plot for Related Schedule (Norm.) Vs Evaluation Criteria

APPENDIX B – COST-QUALIFICATIONS MATRIX AND POST HOC TESTS

Table B.1 Rank and Criteria Combination for Selected Bidders

Combination	Ranking	#1	#2	#3	#4+
Cost/Schedule	#1	25.00%	11.25%	10.00%	3.75%
	#2	8.75%	6.25%	3.75%	3.75%
	#3	6.25%	3.75%	1.25%	2.50%
	#4+	1.25%	2.50%	2.50%	7.50%
Cost/Technical proposals	#1	22.55%	14.71%	3.92%	6.86%
	#2	11.76%	3.92%	1.96%	3.92%
	#3	9.80%	0.00%	0.98%	2.94%
	#4+	9.80%	5.88%	0.00%	0.98%
Cost/PPI	#1	22.55%	11.76%	5.88%	7.84%
	#2	4.90%	4.90%	2.94%	8.82%
	#3	6.86%	1.96%	0.00%	4.90%
	#4+	6.86%	2.94%	2.94%	3.92%
Cost/Interview	#1	20.00%	11.43%	0.00%	0.00%
	#2	14.29%	2.86%	2.86%	0.00%
	#3	17.14%	0.00%	2.86%	0.00%
	#4+	28.57%	0.00%	0.00%	0.00%
Cost/Related experience	#1	49.21%	6.35%	0.00%	0.00%
	#2	19.05%	1.59%	0.00%	0.00%
	#3	9.52%	0.00%	0.00%	0.00%
	#4+	11.11%	1.59%	1.59%	0.00%
Technical proposals/Schedule	#1	17.33%	17.33%	12.00%	9.33%
	#2	2.67%	5.33%	4.00%	5.33%
	#3	12.00%	2.67%	0.00%	1.33%
	#4+	2.67%	1.33%	4.00%	2.67%
Technical proposals/PPI	#1	21.67%	13.33%	5.83%	10.83%
	#2	9.17%	3.33%	5.83%	8.33%
	#3	3.33%	0.83%	0.83%	3.33%
	#4+	5.00%	1.67%	1.67%	5.00%
Technical proposals/Related experience	#1	42.19%	6.25%	0.00%	0.00%
	#2	18.75%	3.13%	1.56%	0.00%
	#3	7.81%	0.00%	0.00%	0.00%
	#4+	20.31%	0.00%	0.00%	0.00%
Interview/Schedule	#1	17.65%	29.41%	5.88%	17.65%
	#2	0.00%	0.00%	5.88%	11.76%
	#3	0.00%	0.00%	5.88%	5.88%
	#4+	0.00%	0.00%	0.00%	0.00%
Interview/Technical proposals	#1	47.17%	20.75%	9.43%	3.77%
	#2	7.55%	5.66%	0.00%	0.00%
	#3	1.89%	1.89%	0.00%	1.89%
	#4+	0.00%	0.00%	0.00%	0.00%
Interview/PPI	#1	26.92%	15.38%	13.46%	25.00%
	#2	1.92%	0.00%	3.85%	7.69%
	#3	1.92%	0.00%	1.92%	1.92%
	#4+	0.00%	0.00%	0.00%	0.00%
Interview/Related experience	#1	53.85%	15.38%	7.69%	0.00%
	#2	7.69%	0.00%	0.00%	0.00%
	#3	15.38%	0.00%	0.00%	0.00%
	#4+	0.00%	0.00%	0.00%	0.00%
PPI/Schedule	#1	16.05%	13.58%	6.17%	2.47%

	#2	9.88%	7.41%	3.70%	1.23%
	#3	3.70%	1.23%	2.47%	1.23%
	#4+	11.11%	1.23%	6.17%	12.35%
PPI/Related experience	#1	35.94%	3.13%	0.00%	0.00%
	#2	20.31%	3.13%	0.00%	0.00%
	#3	7.81%	0.00%	0.00%	0.00%
	#4+	25.00%	3.13%	1.56%	0.00%
Schedule/Related experience	#1	38.33%	5.00%	0.00%	0.00%
	#2	18.33%	3.33%	0.00%	0.00%
	#3	16.67%	1.67%	0.00%	0.00%
	#4+	15.00%	0.00%	1.67%	0.00%

Key: Cost/Interview: Cost on vertical axis, interview on horizontal

Table B.2 Post-hoc testing for Pre-Design Projects

Factor 1	Factor 2	p-value
Related Exp.	PPI	1.000
Related Exp.	Technical Proposals	0.015*
Related Exp.	Interview	0.355
Related Exp.	Cost	0.000*
Related Exp.	Schedule	0.004*
PPI	Technical Proposals	0.674
PPI	Interview	1.000
PPI	Cost	0.008*
PPI	Schedule	0.149
Technical Proposals	Interview	1.000
Technical Proposals	Cost	1.000
Technical Proposals	Schedule	1.000
Interview	Cost	0.836
Interview	Schedule	1.000
Cost	Schedule	1.000

*Statistically Significant at 0.05

Table B.3 Post-hoc testing for Detailed Design Projects

Factor 1	Factor 2	p-value
Related Exp.	PPI	0.832
Related Exp.	Technical Proposals	0.000*
Related Exp.	Interview	0.000*
Related Exp.	Cost	0.000*
Related Exp.	Schedule	0.000*
PPI	Technical Proposals	0.000*
PPI	Interview	0.000*
PPI	Cost	0.000*
PPI	Schedule	0.000*
Technical Proposals	Interview	0.950
Technical Proposals	Cost	0.054
Technical Proposals	Schedule	0.013*
Interview	Cost	1.000
Interview	Schedule	1.000
Cost	Schedule	1.000

*Statistically Significant at 0.05

Table B.4 Post-hoc testing for Architectural Projects

Factor 1	Factor 2	p-value
Related Exp.	PPI	1.000
Related Exp.	Technical Proposals	0.000*
Related Exp.	Interview	0.000*
Related Exp.	Cost	0.000*
Related Exp.	Schedule	0.000*
PPI	Technical Proposals	0.000*
PPI	Interview	0.000*
PPI	Cost	0.000*
PPI	Schedule	0.000*
Technical Proposals	Interview	1.000
Technical Proposals	Cost	0.224
Technical Proposals	Schedule	0.027*
Interview	Cost	1.000
Interview	Schedule	1.000
Cost	Schedule	1.000

*Statistically Significant at 0.05

Table B.5 Post-hoc testing for Engineering Projects

Factor 1	Factor 2	p-value
Related Exp.	PPI	1.000
Related Exp.	Technical Proposals	0.000*
Related Exp.	Interview	0.000*
Related Exp.	Cost	0.000*
Related Exp.	Schedule	0.000*
PPI	Technical Proposals	0.000*
PPI	Interview	0.001*
PPI	Cost	0.000*
PPI	Schedule	0.000*
Technical Proposals	Interview	1.000
Technical Proposals	Cost	0.188
Technical Proposals	Schedule	0.964
Interview	Cost	1.000
Interview	Schedule	1.000
Cost	Schedule	1.000

*Statistically Significant at 0.05

Table B.6 Post-hoc testing for Horizontal Projects

Factor 1	Factor 2	p-value
Related Exp.	PPI	1.000
Related Exp.	Technical Proposals	0.192
Related Exp.	Interview	0.212
Related Exp.	Cost	0.634
Related Exp.	Schedule	0.02*
PPI	Technical Proposals	0.000*
PPI	Interview	0.001*
PPI	Cost	0.04*
PPI	Schedule	0.000*
Technical Proposals	Interview	1.000
Technical Proposals	Cost	1.000
Technical Proposals	Schedule	0.191
Interview	Cost	1.000
Interview	Schedule	0.241
Cost	Schedule	0.137

*Statistically Significant at 0.05

Table B.7 Post-hoc testing for Vertical Projects

Factor 1	Factor 2	p-value
Related Exp.	PPI	0.160
Related Exp.	Technical Proposals	0.000*
Related Exp.	Interview	0.000*
Related Exp.	Cost	0.000*
Related Exp.	Schedule	0.000*
PPI	Technical Proposals	0.000*
PPI	Interview	0.000*
PPI	Cost	0.000*
PPI	Schedule	0.000*
Technical Proposals	Interview	1.000
Technical Proposals	Cost	0.006*
Technical Proposals	Schedule	0.045*
Interview	Cost	1.000
Interview	Schedule	1.000
Cost	Schedule	1.000

*Statistically Significant at 0.05

Table B.8 Post-hoc testing for QBS Projects

Factor 1	Factor 2	p-value
Related Exp.	PPI	-
Related Exp.	Technical Proposals	-
Related Exp.	Interview	-
Related Exp.	Cost	-
Related Exp.	Schedule	-
PPI	Technical Proposals	0.000*
PPI	Interview	0.053
PPI	Cost	-
PPI	Schedule	0.007*
Technical Proposals	Interview	0.498
Technical Proposals	Cost	-
Technical Proposals	Schedule	1.000
Interview	Cost	-
Interview	Schedule	0.207
Cost	Schedule	-

*Statistically Significant at 0.05

Table B.9 Post-hoc testing for BV Projects

Factor 1	Factor 2	p-value
Related Exp.	PPI	0.127
Related Exp.	Technical Proposals	0.000*
Related Exp.	Interview	0.000*
Related Exp.	Cost	0.000*
Related Exp.	Schedule	0.000*
PPI	Technical Proposals	0.000*
PPI	Interview	0.000*
PPI	Cost	0.000*
PPI	Schedule	0.000*
Technical Proposals	Interview	0.054
Technical Proposals	Cost	0.000*
Technical Proposals	Schedule	0.000*
Interview	Cost	1.000
Interview	Schedule	1.000
Cost	Schedule	1.000

*Statistically Significant at 0.05

Table B.10 Post-hoc testing for Small Projects

Factor 1	Factor 2	p-value
Related Exp.	PPI	0.006*
Related Exp.	Technical Proposals	0.000*
Related Exp.	Interview	-
Related Exp.	Cost	0.000*
Related Exp.	Schedule	0.000*
PPI	Technical Proposals	1.000
PPI	Interview	-
PPI	Cost	0.000*
PPI	Schedule	0.000*
Technical Proposals	Interview	-
Technical Proposals	Cost	0.005*
Technical Proposals	Schedule	0.080
Interview	Cost	-
Interview	Schedule	-
Cost	Schedule	1.000

*Statistically Significant at 0.05

Table B.11 Post-hoc testing for Medium Projects

Factor 1	Factor 2	p-value
Related Exp.	PPI	0.626
Related Exp.	Technical Proposals	0.000*
Related Exp.	Interview	0.874
Related Exp.	Cost	0.000*
Related Exp.	Schedule	0.000*
PPI	Technical Proposals	0.064
PPI	Interview	1.000
PPI	Cost	0.000*
PPI	Schedule	0.000*
Technical Proposals	Interview	1.000
Technical Proposals	Cost	1.000
Technical Proposals	Schedule	1.000
Interview	Cost	1.000
Interview	Schedule	1.000
Cost	Schedule	1.000

*Statistically Significant at 0.05

Table B.12 Post-hoc testing for Large Projects

Factor 1	Factor 2	p-value
Related Exp.	PPI	1.000
Related Exp.	Technical Proposals	1.000
Related Exp.	Interview	0.017*
Related Exp.	Cost	1.000
Related Exp.	Schedule	0.051
PPI	Technical Proposals	0.002*
PPI	Interview	0.000*
PPI	Cost	0.002*
PPI	Schedule	0.000*
Technical Proposals	Interview	0.437
Technical Proposals	Cost	1.000
Technical Proposals	Schedule	0.919
Interview	Cost	0.428
Interview	Schedule	1.000
Cost	Schedule	0.905

*Statistically Significant at 0.05

Table B.13 Post-hoc testing for Low Complex Projects

Factor 1	Factor 2	p-value
Related Exp.	PPI	1.000
Related Exp.	Technical Proposals	0.011*
Related Exp.	Interview	0.003*
Related Exp.	Cost	0.002*
Related Exp.	Schedule	0.000*
PPI	Technical Proposals	0.003*
PPI	Interview	0.001*
PPI	Cost	0.000*
PPI	Schedule	0.000*
Technical Proposals	Interview	1.000
Technical Proposals	Cost	1.000
Technical Proposals	Schedule	0.097
Interview	Cost	1.000
Interview	Schedule	1.000
Cost	Schedule	0.909

*Statistically Significant at 0.05

Table B.14 Post-hoc testing for Moderate Complex Projects

Factor 1	Factor 2	p-value
Related Exp.	PPI	0.835
Related Exp.	Technical Proposals	0.000*
Related Exp.	Interview	0.026*
Related Exp.	Cost	0.000*
Related Exp.	Schedule	0.000*
PPI	Technical Proposals	0.044*
PPI	Interview	1.000
PPI	Cost	0.000*
PPI	Schedule	0.000*
Technical Proposals	Interview	1.000
Technical Proposals	Cost	0.114
Technical Proposals	Schedule	1.000
Interview	Cost	0.255
Interview	Schedule	1.000
Cost	Schedule	1.000

*Statistically Significant at 0.05

Table B.15 Post-hoc testing for High Complex Projects

Factor 1	Factor 2	p-value
Related Exp.	PPI	1.000
Related Exp.	Technical Proposals	0.020*
Related Exp.	Interview	0.039*
Related Exp.	Cost	0.000*
Related Exp.	Schedule	0.000*
PPI	Technical Proposals	0.163
PPI	Interview	0.215
PPI	Cost	0.000*
PPI	Schedule	0.000*
Technical Proposals	Interview	1.000
Technical Proposals	Cost	0.694
Technical Proposals	Schedule	0.303
Interview	Cost	1.000
Interview	Schedule	1.000
Cost	Schedule	1.000

*Statistically Significant at 0.05